Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 98 (2007) 706–742
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Statistical inference using higher-order information☆

V.V. Anh[a,*], N.N. Leonenko[b], L.M. Sakhno[c]

[a] *School of Mathematical Sciences, Queensland University of Technology, G.P.O. Box 2434, Brisbane, Qld. 4001, Australia*
[b] *Cardiff School of Mathematics, Cardiff University, Senghennydd Road, Cardiff CF24 4AG, UK*
[c] *Department of Mathematics, Kyiv Taras Shevchenko National University, Volodymyrska 64, 01033 Kyiv, Ukraine*

Received 28 March 2005
Available online 13 November 2006

Abstract

This paper presents a class of minimum contrast estimators for stochastic processes with possible long-range dependence based on the information on higher-order spectral densities. The results on consistency and asymptotic normality of the proposed estimators are provided.

AMS 2000 subject classification: 62M09; 62M15

Keywords: Minimum contrast estimators; Higher-order spectral densities; Long-range dependence; Consistency; Asymptotic normality

1. Introduction

Statistical techniques based on higher-order moments and spectra are of great demand in many fields of applications including geophysics, astronomy, oceanography, sonar, communications, image processing, fluid mechanics, plasma physics, turbulence, economics and finance. The beginning of higher-order statistics can be traced back to Kolmogorov's work and those contributions in the 1960s, such as Brillinger [16] and Brillinger and Rosenblatt [19]; but it is only during recent decades that the area has been rapidly expanding. The bibliography on higher-order statistics

☆ Partially supported by the Australian Research Council Grants DP0345577 and DP0559807, the EPSRC Grant RCMT119 and the NATO Grant PST.CLG.976361.

* Corresponding author. Fax.: +61 7 3864 2310.
E-mail addresses: v.anh@qut.edu.au (V.V. Anh), LeonenkoN@Cardiff.ac.uk (N.N. Leonenko), lms@univ.kiev.ua (L.M. Sakhno).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2006.09.009

compiled by Swami et al. [53] listed 1759 papers, and these papers are only those related to signal processing and engineering applications. Nikias and Petropulu [48] pointed out some motivations behind the use of higher-order spectra in signal processing, namely to detect and characterize non-linearities; to detect signal from Gaussian or non-Gaussian noise; to reconstruct phase and magnitude of signals. In Brillinger [17], a criterion involving the second- and third-order spectral densities was considered with the intention to obtain improved estimates for non-Gaussian time series, as well as for testing the hypothesis of non-Gaussianity. A similar approach was followed in Leonenko et al. [46] (see also [54]).

The second-order information may be identical for different classes of processes. For example, a sequence of independent and identically distributed (iid) random variables and a sequence of uncorrelated non-Gaussian random variables have identical second-order information. To distinguish these two sequences, we have to rely on higher order information. In fact, the third-order spectrum is constant for iid sequences while it can be presented in an expression for uncorrelated non-Gaussian sequences. The second-order information may even be incomplete or useless. For instance, in the classical ARCH-GARCH models, the second-order information is trivial, while the higher orders contain significant information about the parameters of the model. Similar situations occur in some problems of financial mathematics (see [5, Remark 8]). We provide in Section 4 further examples where higher-order information plays an essential role in statistical inference. These examples include Doppler processing using the trispectrum [24,2]; diffusion with linear generator [55,40,39]; bilinear stochastic systems with Brownian motion input [38]; non-Gaussian linear processes [9,10,3]; and non-Gaussian signal in the presence of Gaussian noise. Some more examples are provided in Anh et al. [5,6].

Statistical estimation of random processes and fields in the frequency domain is an extensive area of statistical inference (see, for example, [36,32,33,23,31,50,25,14,28,34,35,29,26,27]). Many of these techniques rely on the information provided by the spectral density, that is, the second-order information only. Kumon [44] treated the identification of a parametric transfer function based on its inner–outer factorization. The outer part is identified by the use of second-order estimates from the observed linear process, while the inner part is identified by the use of higher-order cumulant spectral estimates. Buldygin et al. [20] studied the problem of identification of a linear stochastic system using an integral representation of cumulants of the sample input–output cross-correlogram. Bounds for higher-order cumulants arising from a linear time series regression model were investigated in Zhang and Shaman [56]. The bounds permit derivation of asymptotic moments and asymptotic normality for the model estimators. In connection with the estimation of higher-order spectral densities, we should point out that a non-parametric approach has been mainly followed, resulting in periodograms of higher orders and their modifications with the use of smoothing, tapering, shift-in-time methods (see, for example, [19,1,8]). Estimation of parametric models using higher-order information has been developed in the time domain for some particular cases such as AR and ARMA models based on a generalization of the Yule–Walker method (see, for example, [48]).

In the present paper, we develop a method for minimum contrast estimation based on the spectral densities of the general kth order. The class of processes for which this method is applicable is defined by the conditions on the spectral densities and the weight functions incorporated in the minimum contrast functional, and this approach does not exclude the possibility of long-range dependence in the processes. Furthermore, the corresponding functional is linear with respect to the periodogram of the kth order; therefore there is no need to consider kernel estimation for the spectral densities. To derive the asymptotic properties of our estimators, we use the results on the asymptotic behavior of sample spectral functionals of higher orders. Note that estimates for the spectral functionals of the second order have been extensively studied, but mostly for weakly dependent processes. Results for sample spectral functionals of higher orders can be found, for example, in Bentkus [13] and Keenan [43]; however these results were obtained under some restrictive conditions. Here we provide the results for higher-order spectral functionals under a rather general set of conditions; in fact, we do not impose such conditions as boundedness or square integrability of spectral densities or summability of cumulants. Their proofs are based on the evaluation of the cumulants of the corresponding functionals. Note that cumulants are often used in the derivation of asymptotic probability distributions [30,52] and asymptotic properties of statistical estimators [56,18,20]. We use some ideas on minimum contrast estimation due to Ibragimov [37], Leonenko and Moldavs'ka [45], where analogous functionals based on the second-order spectral density were applied to the estimation of continuous-time random processes and fields with square integrable spectral densities. Anh et al. [4–6] applied these ideas to the case of processes with long-range dependence.

The paper is organized as follows. Section 2 contains the results concerning large sample properties of the sample spectral functionals. These results are then applied to obtain the main results in Section 3 on consistency and asymptotic normality of the proposed minimum contrast estimators. Section 4 provides some examples in which the use of higher-order information is needed for statistical inference. The proofs of the results are grouped together in Section 5. They rely on the properties of the multidimensional kernels of Fejér type and some formulae on cumulants; these are provided in Appendices A and B.

2. Preliminaries

In this section, we present some results on large sample properties of the sample spectral functionals. These results are mostly known for the second-order spectral density [36,11–13]. Their extensions to kth order spectra and corresponding functionals follow the ideas of Brillinger [16], Brillinger and Rosenblatt [19] and Bentkus [13]. Our exposition covers both continuous- and discrete-time stochastic processes. To present the results in a unified manner, we will use Λ to mean either the real line $\mathbb{R}$ or the interval $(-\pi, \pi]$ as appropriate for the continuous- or discrete-time setting, respectively.

Condition I. Let $Y(t)$, $t \in \mathbb{R}$ or $t \in \mathbb{Z}$, be a real-valued measurable strictly stationary process with zero mean and spectral densities $f_k(\lambda_1, \ldots, \lambda_{k-1}) \in L_1\left(\Lambda^{k-1}\right)$ of order $k = 2, 3, \ldots$ such that its cumulant of order k is given by

$$c_k(t_1, \ldots, t_{k-1}) = \int_{\Lambda^{k-1}} f_k(\lambda_1, \ldots, \lambda_{k-1}) e^{i\sum_1^{k-1} \lambda_j t_j} \, d\lambda_1 \ldots d\lambda_{k-1}. \tag{2.1}$$

At some places in the following we will also write the spectral density of the kth order as a function of k variables $f_k(\lambda_1, \ldots, \lambda_k)$, in which case $\lambda_k = -\sum_1^{k-1} \lambda_j$.

Having observed the process $Y(t)$ over the interval $[0, T]$, or given $\{Y(1), \ldots, Y(T)\}$, we define the finite Fourier transform

$$d_T(\lambda) = \int_0^T e^{-i\lambda t} Y(t)\, dt \quad \text{or} \quad d_T(\lambda) = \sum_{t=1}^{T} e^{-it\lambda} Y(\lambda), \quad t \in \Lambda, \tag{2.2}$$

and the periodogram of the kth order

$$I_k^T(\lambda_1, \ldots, \lambda_k) = \frac{1}{(2\pi)^{k-1} T} \prod_{i=1}^{k} d_T(\lambda_i), \tag{2.3}$$

where $\sum_{i=1}^{k} \lambda_i = 0$, but no proper subset of λ_i has sum 0. Consider the spectral functional of the kth order

$$J_k(\varphi) = J_k(\varphi; \psi) = \int_{\Lambda^{k-1}} f_k(\lambda)\, \varphi(\lambda)\, \psi(\lambda)\; d\lambda', \tag{2.4}$$

and the sample spectral functional of the kth order

$$J_k^T(\varphi) = J_k^T(\varphi; \psi) = \int_{\Lambda^{k-1}} I_k^T(\lambda)\, \varphi(\lambda)\, \psi(\lambda)\; d\lambda'. \tag{2.5}$$

In the above integrals, we have denoted $\lambda = (\lambda_1, \ldots, \lambda_k)$ with $\lambda_k = -\sum_{i=1}^{k-1} \lambda_i$, and $\lambda' = (\lambda_1, \ldots, \lambda_{k-1}) \in \Lambda^{k-1}$. The functions $\varphi(\lambda)$ and $\psi(\lambda)$ may be real- or complex-valued and are such that the integral in (2.4) is well defined, and the function $\psi(\lambda)$ satisfies the following condition:

Condition II. $\psi(\lambda) \equiv 0$ if any proper subset of λ_i has sum 0, that is, $\psi(\lambda) \equiv 0$ on all hyperplanes of the form $\sum_{i \in v} \lambda_i = 0$, where $v = \{i_1, \ldots, i_l\} \subset \{1, \ldots, k\}$ and $1 \leqslant l < k$.

Notations: In what follows, we will use the notations similar to (2.4), namely, if it does not cause any confusion, we will write $\int_{\Lambda^{k-1}} g(u)\, du'$, where we will mean $u = (u_1, \ldots, u_k)$ with $u_k = -\sum_{i=1}^{k-1} u_i$ and $u' = (u_1, \ldots, u_{k-1}) \in \Lambda^{k-1}$. Sometimes we will also write such an integral in the form $\int_{\Lambda^{k-1}} g(u) \times \delta\left(\sum_{i=1}^{k} u_i\right) du'$ with $\delta(\cdot)$ being the Kronecker delta function. If v is a set of natural numbers, we will write $|v|$ to denote the number of elements in v, and $\tilde{v}$ to denote the subset of v which contains all the elements of v except the last one. We will also deal with integrals of the form $\int_{\Lambda^{k-p}} g(u) \prod_{l=1}^{p} \delta\left(\sum_{j \in v_l} u_j\right) du'$, where $(v_1, \ldots, v_p)$ is a partition of the set $\{1, \ldots, k\}$. In such a case, integration is understood with respect to $(k-p)$-dimensional vector u', obtained from the vector $u = (u_1, \ldots, u_k)$ in view of p linear restrictions on k variables u_j. For complex-valued variables, we will write $\bar{\xi}$ for the complex conjugate of ξ.

In the statements below, we will use the multidimensional kernels of Fejér type $\Phi_k^T(u)$, $u \in \Lambda^{k-1}$, given in Appendix A by the formulae (A.1) and (A.2) for continuous and discrete time, respectively.

Let us first consider the mean of the functional $J_k^T(\varphi)$.

Lemma 1. *Let the process $Y(t)$, $t \in \mathbb{R}$ or $t \in \mathbb{Z}$, satisfy Condition* I. Then

(1)

$$E J_k^T(\varphi) = \int_{\Lambda^{k-1}} \Phi_k^T(u)\, G_k(u)\; du',$$

where

$$G_k(u) = G_k(u_1, \ldots, u_k; \varphi, \psi)$$
$$= \sum_{\nu=(\nu_1,\ldots,\nu_p)} \int_{\Lambda^{k-p}} \prod_{l=1}^{p} f_{|\nu_l|}\left(\lambda_j + u_j,\, j \in \tilde{\nu}_l\right) H_k(\lambda)$$
$$\times \prod_{l=1}^{p-1} \delta\left(\sum_{j\in\nu_l} \left(\lambda_j + u_j\right)\right) \delta\left(\sum_{i=1}^{k} \lambda_i\right) d\lambda',$$
$$H_k(\lambda) = \varphi(\lambda)\, \psi(\lambda), \tag{2.6}$$

the sum in (2.6) *is taken over all unordered partitions* $(\nu_1, \ldots, \nu_p)$ *of the set* $\{1, \ldots, k\}$.

(2) *If the function* $G_k(u)$ *is bounded and continuous at* $u = 0$, *then as* $T \to \infty$

$$E J_k^T(\varphi) \to \int_{\Lambda^{k-1}} f_k(\lambda)\, H_k(\lambda)\, \delta\left(\sum_{1}^{k} \lambda_i\right) d\lambda'.$$

Lemma 2. *Let the process* $Y(t)$, $t \in \mathbb{R}$ *or* $t \in \mathbb{Z}$, *satisfy Condition* I. *Then*

(1)

$$cov\left(J_k^T\left(\varphi_1\right), J_k^T\left(\varphi_2\right)\right) = \frac{2\pi}{T} \int_{\Lambda^{2k-1}} \Phi_{2k}^T(u)\, G_{2k}(u)\, du',$$

where

$$G_{2k}(u) = G_{2k}\left(u; \varphi_1, \varphi_2, \psi\right)$$
$$= \sum_{\substack{\nu=(\nu_1,\ldots,\nu_p)\\ 1\leqslant p\leqslant k}} \int_{\Lambda^{2k-p-1}} \prod_{l=1}^{p} f_{|\nu_l|}\left(u_j + \lambda_j,\, j \in \tilde{\nu}_l\right)$$
$$\times H_{2k}(\lambda)\, \delta\left(\sum_{j=1}^{k} \lambda_j\right) \delta\left(\sum_{j=k+1}^{2k} \lambda_j\right) \prod_{l=1}^{p-1} \delta\left(\sum_{j\in\nu_l} \left(\lambda_j + u_j\right)\right) d\lambda'. \tag{2.7}$$

Here

$$H_{2k}(\lambda) = \varphi_1(\lambda_1, \ldots, \lambda_k)\, \bar{\varphi}_2(-\lambda_{k+1}, \ldots, -\lambda_{2k})\, \psi(\lambda_1, \ldots, \lambda_k)\, \bar{\psi}(-\lambda_{k+1}, \ldots, -\lambda_{2k})$$

and the sum in (2.7) *is taken over all unordered partitions* $(\nu_1, \ldots, \nu_p)$ *of the table*

1	...	k
$k+1$	...	$2k$

(*see Appendix B*).

(2) *If the function* $G_{2k}(u)$ *is bounded, then*

$$cov\left(J_k^T\left(\varphi_1\right), J_k^T\left(\varphi_2\right)\right) = O\left(\frac{1}{T}\right) \quad as\ T \to \infty.$$

Moreover, if $G_{2k}(u)$ *is bounded and continuous at* $u = 0$, *then as* $T \to \infty$

$$cov\left(T^{1/2} J_k^T\left(\varphi_1\right), T^{1/2} J_k^T\left(\varphi_2\right)\right) = 2\pi G_{2k}(0).$$

We have the following consequence of Lemmas 1 and 2.

Lemma 3. *Let the conditions of Lemmas* 1 *and* 2 *be satisfied. Then* $J_k^T(\varphi) \to J_k(\varphi)$ *in probability as* $T \to \infty$.

Lemma 4. *Let the process* $Y(t)$, $t \in \mathbb{R}$ *or* $t \in \mathbb{Z}$, *satisfy Condition* I *and* $\left\{\varphi_1, \ldots, \varphi_m\right\}$, $m \geqslant 2$, *be a set of weight functions. Then*

(1)

$$cum\left(J_k^T\left(\varphi_1\right), \ldots, J_k^T\left(\varphi_m\right)\right) = \frac{(2\pi)^{m-1}}{T^{m-1}} \int_{\Lambda^{km-1}} \Phi_{km}^T(u)\, G_{km}(u)\, du',$$

where

$$\begin{aligned} G_{km}(u) &= G_{km}\left(u; \varphi_1, \ldots, \varphi_m, \psi\right) \\ &= \sum_{v=(v_1,\ldots,v_p)} \int_{\Lambda^{(m-1)k-p+1}} \prod_{i=1}^{m} \varphi_i\left(\lambda_{(i-1)k+1}, \ldots, \lambda_{ik}\right) \\ &\quad \times \prod_{i=1}^{m} \psi\left(\lambda_{(i-1)k+1}, \ldots, \lambda_{ik}\right) \prod_{i=1}^{p} f_{|v_i|}\left(u_j + \lambda_j, j \in \tilde{v}_i\right) \\ &\quad \times \prod_{l=1}^{p-1} \delta\left(\sum_{j \in v_l}\left(u_j + \lambda_j\right)\right) \prod_{i=1}^{m} \delta\left(\lambda_{(i-1)k+1} + \cdots + \lambda_{ik}\right) d\lambda'. \end{aligned} \tag{2.8}$$

The sum in (2.8) *is taken over all indecomposable partitions* $\left(v_1, \ldots, v_p\right)$ *of the table*

1	$\ldots$	k
$k+1$	$\ldots$	$2k$
$\ldots$	$\ldots$	$\ldots$
$m(k-1)+1$	$\ldots$	mk

(2) *If the function* $G_{km}\left(u; \varphi_1, \ldots, \varphi_m, \psi\right)$ *is bounded, then*

$$cum\left(J_k^T\left(\varphi_1\right), \ldots, J_k^T\left(\varphi_m\right)\right) = O\left(\frac{1}{T^{m-1}}\right) \quad \text{as } T \to \infty.$$

We are able to present more general formulae for the cumulants of the functionals $J_k^T(\varphi)$. Consider the cumulant of order $p \geqslant 2$:

$$cum\left(J_{k_1}^T\left(g_{k_1}\right), \ldots, J_{k_p}^T\left(g_{k_p}\right)\right) = cum\left(J_{k_1}^T\left(g_{k_1}, \psi_{k_1}\right), \ldots, J_{k_p}^T\left(g_{k_p}, \psi_{k_p}\right)\right)$$

for some fixed set of indices $\{k_1, \ldots, k_p\}$, $k_i \geqslant 2$, $i = 1, \ldots, p$, a set of weight functions $\{g_{k_i}\}_{i=1,\ldots p} = \{g_{k_i}(\lambda_1, \ldots, \lambda_{k_i})\}_{i=1,\ldots p}$ and the corresponding set of functions $\{\psi_{k_i}\}_{i=1,\ldots p} =$

$\{\psi_{k_i}(\lambda_1, \ldots, \lambda_{k_i})\}_{i=1,\ldots p}$, which satisfy Condition II. Let us denote $N_q = k_1 + \cdots + k_q$, $q = 1, \ldots, p$, $N = N_p$, $N_0 = 0$.

Lemma 5. *Let the process $Y(t)$, $t \in \mathbb{R}$ or $t \in \mathbb{Z}$, satisfy Condition* I. *Then*:

(1)

$$cum\left(J_{k_1}^T\left(g_{k_1}\right), \ldots, J_{k_p}^T\left(g_{k_p}\right)\right) = \frac{(2\pi)^{N-1}}{T^{N-1}} \int_{\Lambda^{N-1}} \Phi_N^T(u)\, F_{k_1\ldots k_p}(u)\, du',$$

where

$$\begin{aligned} F_{k_1\ldots k_p}(u) = F_{k_1\ldots k_p}&\left(u; g_{k_1}, \ldots, g_{k_p}, \psi_{k_1}, \ldots, \psi_{k_p}\right) \\ &= \sum_{v=(v_1,\ldots,v_l)} \int_{\Lambda^{N-l-p+1}} \prod_{i=1}^{p} g_{k_i}\left(\lambda_{N_{i-1}+1}, \ldots, \lambda_{N_i}\right) \\ &\quad \times \prod_{i=1}^{p} \psi_{k_i}\left(\lambda_{N_{i-1}+1}, \ldots, \lambda_{N_i}\right) \prod_{i=1}^{l} f_{|v_i|}\left(u_j + \lambda_j, j \in \tilde{v}_i\right) \\ &\quad \times \prod_{i=1}^{l} \delta\left(\sum_{j \in v_i}\left(u_j + \lambda_j\right)\right) \prod_{i=1}^{p-1} \delta\left(\sum_{j=N_{i-1}}^{N_i} \lambda_j\right) d\lambda'. \end{aligned} \tag{2.9}$$

The sum in (2.9) *is taken over all indecomposable partitions* $(v_1, \ldots, v_l)$ *of the table*

1	…	N_1
$N_1 + 1$	…	N_2
…	…	…
$N_{p-1} + 1$	…	N_p

(*see Appendix B*).

(2) *If the function $F_{k_1\ldots k_p}(u)$ is bounded, then*

$$cum\left(J_{k_1}^T\left(g_{k_1}\right), \ldots, J_{k_p}^T\left(g_{k_p}\right)\right) = O\left(\frac{1}{T^{N-1}}\right) \quad as\ T \to \infty.$$

Now let us fix the weight functions $\varphi_1, \ldots, \varphi_m$ and set

$$J_k^T = \left\{J_k^T\left(\varphi_i\right)\right\}_{i=1,\ldots,m} = \left\{\int_{\Lambda^{k-1}} I_k^T(\lambda)\, \varphi_i(\lambda)\, \psi(\lambda)\, d\lambda'\right\}_{i=1,\ldots,m}$$

and

$$J_k = \left\{J_k\left(\varphi_i\right)\right\}_{i=1,\ldots,m} = \left\{\int_{\Lambda^{k-1}} f_k(\lambda)\, \varphi_i(\lambda)\, \psi(\lambda)\, d\lambda'\right\}_{i=1,\ldots,m}.$$

Let $\xi = \{\xi_i\}_{i=1,\ldots,m}$ be a complex-valued Gaussian random vector with mean zero and second-order moments

$$w_{ij} = E\xi_i\bar{\xi}_j = 2\pi G_{2k}\left(0; \varphi_i, \varphi_j, \psi\right), \quad i, j = 1, \ldots, m,$$

where $G_{2k}\left(u; \varphi_i, \varphi_j, \psi\right)$ is given by (2.7).

Condition III. The weight functions $\varphi_1, \ldots, \varphi_m$ and the spectral density of the kth order f_k are such that

$$T^{1/2}\left(EJ_k^T - J_k\right) \to 0 \quad \text{as } T \to \infty.$$

Lemma 6. *Let the assumptions of Lemmas* 1 *and* 2 *hold and the functions* $G_{kl}\left(u; \varphi_{m_1}, \ldots, \varphi_{m_l}; \psi\right)$ *defined by* (2.8) *are bounded for all* $l = 2, 3, \ldots$ *and all choices* $(m_1, \ldots, m_l)$ *with* $1 \leqslant m_i \leqslant m$, $i = 1, \ldots, l$. *Then as* $T \to \infty$

$$T^{1/2}\left(J_k^T - EJ_k^T\right) \xrightarrow{d} \xi \tag{2.10}$$

and, moreover, if the assumption III *holds, then as* $T \to \infty$

$$T^{1/2}\left(J_k^T - J_k\right) \xrightarrow{d} \xi. \tag{2.11}$$

3. Main results

Let a random process $Y(t)$ be observed on the interval $[0, T]$ or at the time points $t = 1, 2, \ldots, T$. Consider a parametric statistical model with a family of distributions $\{P_\theta, \theta \in \Theta\}$, where Θ is a compact subset of $\mathbb{R}^m$ and the true value of the parameter vector $\theta_0 \in int\, \Theta$, the interior of Θ. Denote $P_0 = P_{\theta_0}$.

A non-random real-valued function $K(\theta_0; \theta) \geqslant 0$ is called a *contrast function* if it has a unique minimum at $\theta = \theta_0$. A random process $U_T(\theta)$, $T \in \mathbb{R}$ or $T \in \mathbb{N}$, $\theta \in \Theta$, related to the observations $\{Y(t), t \in [0, T]\}$ or $\{Y(t), t = 1, 2, \ldots, T\}$ is called the *contrast process* for a contrast function $K(\theta_0; \theta)$ if it satisfies the inequality

$$\liminf_{T \to \infty} \left[U_T(\theta) - U_T(\theta_0)\right] \geqslant K(\theta_0; \theta)$$

in P_0-probability. The *minimum contrast estimator* $\hat{\theta}_T$ is defined as a minimum point of the functional $U_T(\theta)$, that is,

$$\hat{\theta}_T = \arg\min_{\theta \in \Theta} U_T(\theta).$$

We will use the condition of *identifiability* of a random process in the following sense: having introduced a parametric family of probability measures $\{P_\theta, \theta \in \Theta\}$, we will suppose that (i) different values of the parameter vector θ correspond to different probability measures P_θ, and each probability measure P_θ in turn, in accordance with Kolmogorov's consistency theorem, corresponds to a certain random process possessing a particular family of spectral densities (induced by this probability measure P_θ), and (ii) for $\theta_1, \theta_2 \in \Theta$ with $\theta_1 \neq \theta_2$, we will have for the spectral densities of order $k = 2, 3, \ldots$

$$f_k(\lambda_1, \ldots, \lambda_{k-1}, \theta_1) \neq f_k(\lambda_1, \ldots, \lambda_{k-1}, \theta_2)$$

almost everywhere in Λ^{k-1} with respect to Lebesgue measure.

In what follows, we will always assume that the above condition of identifiability is satisfied. Let us introduce the following conditions.

Condition IV. Let the random process $Y(t)$, $t \in \mathbb{R}$ or $t \in \mathbb{Z}$, satisfy Condition I and suppose that its spectral densities depend on an unknown parameter θ:

$$f_k(\lambda) = f_k(\lambda_1, \ldots, \lambda_{k-1}; \theta) = \operatorname{Re} f_k(\lambda; \theta) + i \operatorname{Im} f_k(\lambda; \theta)$$
$$= f_k^{(1)}(\lambda; \theta) + i f_k^{(2)}(\lambda; \theta), \quad \lambda = (\lambda_1, \ldots, \lambda_{k-1}) \in \Lambda^{k-1},$$

where $\theta \in \Theta \subset \mathbb{R}^m$, Θ being a compact set, and the true value of the parameter $\theta_0 \in \mathit{int}\,\Theta$.

Condition V. Let the real-valued functions $w_k^{(i)}(\lambda)$, $i = 1, 2$, $w_{k,0}(\lambda)$ and the spectral density of the kth order satisfy the following assumptions:

(i) $w_k^{(i)}(\lambda)$, $i = 1, 2$, and $w_{k,0}(\lambda)$ satisfy the same conditions of symmetry as the kth order spectral density;
(ii) $w_{k,0}(\lambda)$ satisfies Condition II;
(iii) $w_k^{(i)}(\lambda)\, w_{k,0}(\lambda)\, f_k^{(i)}(\lambda; \theta) \in L_1\left(\Lambda^{k-1}\right)$, $i = 1, 2$, for all $\theta \in \Theta$;
(iv) $w_k^{(i)}(\lambda)\, f_k^{(i)}(\lambda; \theta) \geqslant 0$, $i = 1, 2$, $(\lambda; \theta) \in \Lambda^{k-1} \times \Theta$.

Under Condition V we set

$$\int_{\Lambda^{k-1}} f_k^{(i)}(\lambda; \theta)\, w_k^{(i)}(\lambda)\, w_{k,0}(\lambda)\, d\lambda' = \sigma_k^{(i)}(\theta), \quad i = 1, 2 \tag{3.1}$$

and consider the factorization of the real and imaginary parts of the spectral density $f_k(\lambda; \theta)$:

$$f_k^{(i)}(\lambda; \theta)\, w_k^{(i)}(\lambda) = \sigma_k^{(i)}(\theta)\, \psi_k^{(i)}(\lambda; \theta), \quad i = 1, 2, \ \theta \in \Theta,$$
$$\lambda = (\lambda_1, \ldots, \lambda_k), \quad \lambda_k = -\sum_{j=1}^{k-1} \lambda_j, \ \ \lambda_j \in \Lambda, \ \ j = 1, \ldots, k. \tag{3.2}$$

The functions $\psi_k^{(i)}(\lambda; \theta)$, $i = 1, 2$, satisfy

$$\psi_k^{(i)}(\lambda; \theta) = \frac{f_k^{(i)}(\lambda; \theta)\, w_k^{(i)}(\lambda)}{\sigma_k^{(i)}(\theta)} \tag{3.3}$$

and

$$\int_{\Lambda^{k-1}} \psi_k^{(i)}(\lambda; \theta)\, w_{k,0}(\lambda)\, d\lambda' = 1. \tag{3.4}$$

In the following we will omit the index k in the functions $w_k^{(i)}$ and $w_{k,0}$ if this does not cause any confusion.

Consider the following contrast process:

$$U_T(\theta) = -\Bigg(p \int_{\Lambda^{k-1}} \operatorname{Re} I_k^T(\lambda)\, w^{(1)}(\lambda)\, w_0(\lambda) \log \psi_k^{(1)}(\lambda; \theta)\, d\lambda'$$
$$+ q \int_{\Lambda^{k-1}} \operatorname{Im} I_k^T(\lambda)\, w^{(2)}(\lambda)\, w_0(\lambda) \log \psi_k^{(2)}(\lambda; \theta)\, d\lambda'\Bigg), \tag{3.5}$$

where $I_k^T(\lambda)$ is the periodogram of the kth order given by (2.3), the non-negative numbers p and q are such that $p+q=1$.

We introduce the functions

$$\mathcal{K}(\theta_0;\theta) = p\int_{\Lambda^{k-1}} f_k^{(1)}(\lambda;\theta_0)\log\frac{\psi_k^{(1)}(\lambda;\theta_0)}{\psi_k^{(1)}(\lambda;\theta)} w^{(1)}(\lambda)\, w_0(\lambda)\, d\lambda' + q\int_{\Lambda^{k-1}} f_k^{(2)}(\lambda;\theta_0)\log\frac{\psi_k^{(2)}(\lambda;\theta_0)}{\psi_k^{(2)}(\lambda;\theta)} w^{(2)}(\lambda)\, w_0(\lambda)\, d\lambda' \tag{3.6}$$

and

$$U(\theta) = -\left(p\int_{\Lambda^{k-1}} f_k^{(1)}(\lambda;\theta_0)\, w^{(1)}(\lambda)\, w_0(\lambda)\log\psi_k^{(1)}(\lambda;\theta)\, d\lambda' + q\int_{\Lambda^{k-1}} f_k^{(2)}(\lambda;\theta_0)\, w^{(2)}(\lambda)\, w_0(\lambda)\log\psi_k^{(2)}(\lambda;\theta)\, d\lambda'\right). \tag{3.7}$$

Remark 1. The aim behind the introduction of the weight function $w(\lambda)$ is twofold:

(i) This function compensates possible singularities of the spectral density; in other words, it is needed to control the behavior of the integrand in the spectral functional (3.5) (and related integrals) at the points of singularity.
(ii) Dealing with continuous-time random processes and continuous observations, a weight function is needed to warrant the convergence of corresponding integrals.

Obviously, we cannot select a weight function $w(\lambda)$ to suit all density functions. This selection depends on the specific form of the spectral density. Furthermore, the two aspects (i) and (ii) induce a set of conditions needed for the function $w(\lambda)$ formulated below. A contrast process of the type (3.5) was used in Leonenko and Moldavs'ka [45] where, for the estimation of random fields with spectral density $f(\lambda;\theta)\in L_2(\mathbb{R}^n)$, the weight function of the form $w(\lambda)=\frac{1}{1+|\lambda|^2}$ was chosen. The same functional (3.5) with a specific weight function was used in Anh et al. [4] for parameter estimation of fractional Riesz–Bessel motion (FRBM). The idea to introduce a weight function into a contrast process for the case of continuous-time stochastic processes (observed continuously) was also used in Gao et al. [27]. In their paper, a continuous version of the Gauss–Whittle contrast function (with weight function $w(\lambda)=\frac{1}{1+\lambda^2}$) was considered for the estimation of FRBM.

Remark 2. For the case of second-order spectral density, we suppose the existence of a real-valued, non-negative function $w(\lambda)$, $\lambda\in\Lambda$, symmetric about zero, such that $w(\lambda)\, f(\lambda;\theta)\in L_1(\Lambda)$, $\forall\theta\in\Theta$, and introduce the factorization

$$f_2(\lambda;\theta) = \sigma^2(\theta)\,\psi(\lambda;\theta), \quad \lambda\in\Lambda,\ \theta\in\Theta,$$

where

$$\sigma^2(\theta) = \int_\Lambda f_2(\lambda;\theta)\, w(\lambda)\, d\lambda$$

and

$$\int_{\Lambda} \psi(\lambda;\theta)\, w(\lambda)\, d\lambda = 1.$$

The contrast field in this case is of the form

$$U_{2,T}(\theta) = -\int_{\Lambda} I_{2,T}(\lambda)\, w(\lambda) \log \psi(\lambda;\theta)\, d\lambda,$$

where $I_{2,T}(\lambda)$ is a periodogram of the second order given by (2.3) with $k = 2$. This case was treated in Anh et al. [4].

Remark 3. Continuing our comment regarding weight functions, let us consider again the case $k = 2$ and the functional $U_{2,T}(\theta)$. We note the following fact: if there exists a smooth function $v : \Lambda \to \Lambda$ with Fourier transform $\hat{v}$ such that $w(\lambda) = |\hat{v}(\lambda)|^2$, then the product $f_2(\lambda)w(\lambda)$ can be viewed as the spectral density of the random process obtained from the original one by linear filtering with transfer function $\hat{v}(\lambda)$ (i.e., impulse response function $v(\lambda)$). Here we find a parallel situation with Heyde and Gay [34] where the asymptotics for the smoothed periodogram were derived via filtering of the original process, which may not have a square integrable spectral density, to produce a related one for which the spectral density is square integrable. Some standard results on the asymptotics of corresponding covariances can then be applied. The set of assumptions to be satisfied by the spectral density and smoothing function was tailored to implement this idea.

We will need the following conditions:

Condition VI. The derivatives $\nabla_\theta \psi_k^{(i)}(\lambda;\theta)$, $i = 1, 2$, exist and

$$\nabla_\theta \int_{\Lambda^{k-1}} \psi_k^{(i)}(\lambda;\theta)\, w_0(\lambda)\, d\lambda' = \int_{\Lambda^{k-1}} \nabla_\theta \psi_k^{(i)}(\lambda;\theta)\, w_0(\lambda)\, d\lambda' = 0, \quad i = 1, 2. \tag{3.8}$$

Condition VII. The functions $G_k\left(u; w^{(i)} \log \psi_k^{(i)}, w_0\right)$, $i = 1, 2$, are bounded and continuous at the point $u = 0$, the functions $G_k\left(u; w^{(i)}, w_0\right)$, $i = 1, 2$, are bounded ($G_k(u; \varphi, \psi)$ is defined by (2.6)).

Condition VIII. The functions $G_{2k}\left(u; w^{(i)} \log \psi_k^{(i)}, w^{(i)} \log \psi_k^{(i)}, w_0\right)$, $i = 1, 2$,

(i) are bounded;
(ii) are continuous at the point $u = 0$, where the function $G_{2k}\left(u; \varphi_1, \varphi_2, \psi\right)$ is defined by (2.7).

Condition IX. There exist the functions $v_1(\lambda)$ and $v_2(\lambda)$ such that

(i) the functions

$$h_k^{(i)}(\lambda;\theta) = v_i(\lambda) \log \psi_k^{(i)}(\lambda;\theta), \quad i = 1, 2$$

are uniformly continuous in $\Lambda^{k-1} \times \Theta$;

(ii) the functions $G_k\left(u; \frac{w^{(i)}}{v_i}, w_0\right)$, $i = 1, 2$, are bounded and continuous at $u = 0$ and the functions $G_{2k}\left(u; \frac{w^{(i)}}{v_i}, \frac{w^{(i)}}{v_i}, w_0\right)$, $i = 1, 2$, are bounded.

Condition X. $w_k^{(1)}(\lambda)\operatorname{Re} I_k^T(\lambda) \geqslant 0$, $w_k^{(2)}(\lambda)\operatorname{Im} I_k^T(\lambda) \geqslant 0$.

Theorem 1. *Let the random process $Y(t)$, $t \in \mathbb{R}$ or $t \in \mathbb{Z}$, satisfy Condition* IV, *and Conditions* V, VII, VIII(i), IX *and* X *hold*.

Then the function $\mathcal{K}(\theta_0;\theta)$ defined by (3.6) *is the contrast function for the contrast process $U_T(\theta)$ defined by* (3.5). *The minimum contrast estimator $\hat{\theta}_T$ defined as*

$$\hat{\theta}_T = \operatorname*{arg\,min}_{\theta\in\Theta} U_T(\theta) \tag{3.9}$$

is a consistent estimator of the parameter θ_0, that is, $\hat{\theta}_T \to \theta_0$ in *P_0-probability as $T\to\infty$, and the estimators*

$$\hat{\sigma}_{k,T}^{(1)} = \int_{\Lambda^{k-1}} \operatorname{Re} I_k^T(\lambda)\, w^{(1)}(\lambda)\, w_0(\lambda)\, d\lambda' \tag{3.10}$$

and

$$\hat{\sigma}_{k,T}^{(2)} = \int_{\Lambda^{k-1}} \operatorname{Im} I_k^T(\lambda)\, w^{(2)}(\lambda)\, w_0(\lambda)\, d\lambda' \tag{3.11}$$

are consistent estimators of $\sigma_k^{(1)}(\theta_0)$ and $\sigma_k^{(2)}(\theta_0)$, respectively.

To state the result on the asymptotic normality of the minimum contrast estimator (3.9), we will need some additional conditions.

Condition XI. The functions $\psi_k^{(i)}(\lambda;\theta)$, $i=1,2$, are twice differentiable in the neighborhood of the point θ_0 and the functions

$$\varphi_l^{ij}(\lambda;\theta) = w^{(l)}(\lambda)\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\psi_k^{(l)}(\lambda;\theta), \quad i,j=1,\ldots,m,\ l=1,2,\ \theta\in\Theta \tag{3.12}$$

and

$$g_k^{(i)}(\lambda;\theta) = w^{(1)}(\lambda)\frac{\partial}{\partial\theta_i}\log\psi_k^{(1)}(\lambda;\theta), \quad i=1,\ldots,m,\ \theta\in\Theta,$$

$$g_k^{(i+m)}(\lambda;\theta) = w^{(2)}(\lambda)\frac{\partial}{\partial\theta_i}\log\psi_k^{(2)}(\lambda;\theta), \quad i=1,\ldots,m,\ \theta\in\Theta \tag{3.13}$$

are such that

(i) the functions $G_k(u;\varphi_l^{ij},w_0)$, $i,j=1,\ldots,m$, $l=1,2$, are bounded and continuous at $u=0$ for all $\theta\in\Theta$;
(ii) the functions $G_{2k}(u;\varphi_l^{ij},\varphi_l^{ij},w_0)$, $i,j=1,\ldots,m$, $l=1,2$, are bounded for all $\theta\in\Theta$;
(iii) the functions $G_{kl}(u;g_k^{(m_1)},\ldots,g_k^{(m_l)},w_0)$, defined by (2.8) are bounded for all $\theta\in\Theta$, $l=2,3,\ldots$ and all choices $(m_1,\ldots,m_l)$, $1\leqslant m_i\leqslant 2m$, $i=1,\ldots,l$;
(iv) Condition III holds for $\left\{J_k^T\left(g_k^{(i)};w_0\right)\right\}_{i=1,\ldots,2m}$ for all $\theta\in\Theta$;
(v) the second-order derivatives $\left(\partial^2/\partial\theta_i\partial\theta_j\right)\log\psi_k^{(l)}(\lambda;\theta)$, $i,j=1,\ldots,m$, $l=1,2$, are continuous in θ.

Condition XII. The matrices $S_k(\theta) = \left\{ s_{ij}^{(k)}(\theta) \right\}_{i,j=1,\dots,m}$ and $A_k(\theta) = \left\{ a_{ij}^{(k)}(\theta) \right\}_{i,j=1,\dots,m}$ are positive definite, where

$$\begin{aligned} s_{ij}^{(k)}(\theta) &= p \int_{\Lambda^{k-1}} f_k^{(1)}(\lambda;\theta)\,\varphi_1^{ij}(\lambda;\theta)\,w_0(\lambda)\,d\lambda' \\ &\quad + q \int_{\Lambda^{k-1}} f_k^{(2)}(\lambda;\theta)\,\varphi_2^{ij}(\lambda;\theta)\,w_0(\lambda)\,d\lambda' \\ &= \sigma_k^{(1)}(\theta)\, p \int_{\Lambda^{k-1}} \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j}\psi_k^{(1)} - \frac{1}{\psi_k^{(1)}} \frac{\partial}{\partial\theta_i}\psi_k^{(1)} \frac{\partial}{\partial\theta_j}\psi_k^{(1)} \right) d\lambda' \\ &\quad + \sigma_k^{(2)}(\theta)\, q \int_{\Lambda^{k-1}} \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j}\psi_k^{(2)} - \frac{1}{\psi_k^{(2)}} \frac{\partial}{\partial\theta_i}\psi_k^{(2)} \frac{\partial}{\partial\theta_j}\psi_k^{(2)} \right) d\lambda', \\ &\quad i,j = 1,\dots,m, \end{aligned} \tag{3.14}$$

$$\begin{aligned} a_{ij}^{(k)}(\theta) &= \tfrac{1}{2} \Big\{ p^2 \operatorname{Re} G_{2k}\left(0; g_k^{(i)}, g_k^{(j)}, w_0\right) + q^2 \operatorname{Re} G_{2k}\left(0; g_k^{(i+m)}, g_k^{(j+m)}, w_0\right) \\ &\quad + pq \operatorname{Im} G_{2k}\left(0; g_k^{(i+m)}, g_k^{(j)}, w_0\right) - pq \operatorname{Im} G_{2k}\left(0; g_k^{(i)}, g_k^{(j+m)}, w_0\right) \Big\}, \\ &\quad i,j = 1,\dots,m, \end{aligned} \tag{3.15}$$

where $G_{2k}\left(u;\varphi_1,\varphi_2,w_0\right)$ is given by (2.7).

Theorem 2. *Let Conditions* IV–XII *be satisfied. Then as* $T \to \infty$

$$T^{1/2}\left(\widehat{\theta}_T - \theta_0\right) \xrightarrow{d} N_m\left(0, S_k^{-1}(\theta_0)\,A_k(\theta_0)\,S_k^{-1}(\theta_0)\right), \tag{3.16}$$

where the matrices $S_k(\theta)$ *and* $A_k(\theta)$ *are defined in Condition* XII.

We can also construct the contrast process involving two spectral densities of the order k_1 and k_2. Let the spectral densities f_{k_1} and f_{k_2} satisfy Condition V with weight functions $w_{k_1}^{(i)}, w_{k_1,0}$ and $w_{k_2}^{(i)}, w_{k_2,0}$, respectively, and assume that these spectral densities can be factorized in the form (3.2).

Consider the following contrast process

$$\begin{aligned} V_T(\theta) = -\sum_{k=k_1,k_2} \Bigg(p_k \int_{\Lambda^{k-1}} \operatorname{Re} I_k^T(\lambda)\,w_k^{(1)}(\lambda)\,w_{k,0}(\lambda) \log\psi_k^{(1)}(\lambda;\theta)\,d\lambda' \\ + q_k \int_{\Lambda^{k-1}} \operatorname{Im} I_k^T(\lambda)\,w_k^{(2)}(\lambda)\,w_{k,0}(\lambda) \log\psi_k^{(2)}(\lambda;\theta)\,d\lambda' \Bigg), \end{aligned} \tag{3.17}$$

where $\sum_{k=k_1,k_2}(p_k+q_k) = 1$. Here, the argument λ is understood as either k_1- or k_2-dimensional in the corresponding integrals, in accordance with our system of notations introduced in Section 2.

Similar to the above, we introduce the function

$$\begin{aligned} \mathcal{K}_V(\theta_0;\theta) = \sum_{k=k_1,k_2} \Bigg(p_k \int_{\Lambda^{k-1}} f_k^{(1)}(\lambda;\theta_0) \log\frac{\psi_k^{(1)}(\lambda;\theta_0)}{\psi_k^{(1)}(\lambda;\theta)}\,w_k^{(1)}(\lambda)\,w_{k,0}(\lambda)\,d\lambda' \\ + q_k \int_{\Lambda^{k-1}} f_k^{(2)}(\lambda;\theta_0) \log\frac{\psi_k^{(2)}(\lambda;\theta_0)}{\psi_k^{(2)}(\lambda;\theta)}\,w_k^{(2)}(\lambda)\,w_{k,0}(\lambda)\,d\lambda' \Bigg). \end{aligned} \tag{3.18}$$

Theorem 3. *Let the assumptions of Theorem* 1 *be satisfied for both spectral densities* f_{k_1} *and* f_{k_2} *and corresponding weight functions* $w_{k_i}^{(l)}, w_{k_i,0}, i = 1, 2, l = 1, 2$. *Then the function* $\mathcal{K}_V(\theta_0; \theta)$ *defined by* (3.17) *is the contrast function for the contrast process* $V_T(\theta)$ *defined by* (3.16). *The minimum contrast estimator* $\hat{\theta}_T^V$ *defined as*

$$\hat{\theta}_T^V = \arg\min_{\theta\in\Theta} V_T(\theta) \tag{3.19}$$

is a consistent estimator of the parameter θ, *that is,* $\hat{\theta}_T \to \theta_0$ *in* P_0*-probability as* $T \to \infty$, *and the estimators* $\hat{\sigma}_{k,T}^{(1)}$ *and* $\hat{\sigma}_{k,T}^{(2)}$, $k = k_1, k_2$, *given by the formulae* (3.10) *and* (3.11) *are consistent estimators of* $\sigma_k^{(1)}(\theta)$ *and* $\sigma_k^{(2)}(\theta)$, $k = k_1, k_2$, *respectively*.

To state the asymptotic normality of the minimum contrast estimator (3.18), we will need two more conditions in addition to all the above conditions of this section.

Condition XIII. The functions $F_{q_1\ldots q_p}(u) = F_{q_1\ldots q_p}\left(u; g_{q_1}^{(i_1)}, \ldots, g_{q_p}^{(i_p)}, w_{q_1,0}, \ldots, w_{q_p,0}\right)$ given by (2.9) are bounded for all $p \geqslant 2$, all choices $(q_1, \ldots, q_p)$, $q_i \in \{k_1, k_2\}$, and all choices of the set of weight functions $\{g_{q_1}^{(i_1)}, \ldots, g_{q_p}^{(i_p)}\}$ and corresponding set of functions $\{w_{q_1,0}, \ldots, w_{q_p,0}\}$, $1 \leqslant i_j \leqslant 2m$, $q_i \in \{k_1, k_2\}$, for all $\theta \in \Theta$ (the functions $g_k^{(i)}$ are given by the formula (3.13)).

Condition XIV. The matrix $A_{k_1k_2}(\theta) = \left\{a_{ij}^{(k_1k_2)}(\theta)\right\}_{i,j=1,\ldots,m}$ is positive definite, where

$$\begin{aligned} a_{ij}^{(k_1k_2)}(\theta) = a_{ij}^{(k_1)}(\theta) + a_{ij}^{(k_2)}(\theta) & \\ & + p_{k_1} p_{k_2} \operatorname{Re} F_{k_1k_2}\left(0; g_{k_1}^{(i)}, g_{k_2}^{(j)}, w_{k_1,0}, w_{k_2,0}\right) \\ & + q_{k_1} q_{k_2} \operatorname{Re} F_{k_1k_2}\left(0; g_{k_1}^{(i+m)}, g_{k_2}^{(j+m)}, w_{k_1,0}, w_{k_2,0}\right) \\ & - p_{k_1} q_{k_2} \operatorname{Im} F_{k_1k_2}\left(0; g_{k_1}^{(i)}, g_{k_2}^{(j+m)}, w_{k_1,0}, w_{k_2,0}\right) \\ & + p_{k_2} q_{k_1} \operatorname{Im} F_{k_1k_2}\left(0; g_{k_1}^{(i+m)}, g_{k_2}^{(j)}, w_{k_1,0}, w_{k_2,0}\right), \\ & i, j = 1, \ldots, m. \end{aligned} \tag{3.20}$$

Theorem 4. *Let Conditions* IV–XII *with* $k = k_1$ *and* $k = k_2$ *and Conditions* XIII–XIV *be satisfied. Then as* $T \to \infty$

$$T^{1/2}\left(\widehat{\theta}_T^V - \theta_0\right) \xrightarrow{d} N_m\left(0, S_{k_1k_2}^{-1}(\theta_0) A_{k_1k_2}(\theta_0) S_{k_1k_2}^{-1}(\theta_0)\right), \tag{3.21}$$

where $S_{k_1k_2}(\theta) = S_{k_1}(\theta) + S_{k_2}(\theta)$, *the matrices* $S_k(\theta)$ *and* $A_{k_1k_2}(\theta)$ *being defined by* (3.14) *and* (3.20), *respectively*.

Remark 4. We should point out one particular situation where the proposed estimation method is not quite complete, and some additional steps need to be taken. Namely, let us consider the models with spectral densities $f_k(\lambda_1, \ldots, \lambda_{k-1})$, $k \geqslant 2$, of the following form:

$$f_k\left(\lambda_1, \ldots, \lambda_{k-1}; \theta'\right) = f_k(\lambda_1, \ldots, \lambda_{k-1}; \eta, \theta) = h_k(\eta)\, g_k(\lambda_1, \ldots, \lambda_{k-1}; \theta),$$

that is, the parameter vector θ' consists of two parts, $\theta' = (\eta, \theta)$, and the (sub)vector η appears only in the multiplicative term in the expressions for the spectral densities (see, for example, the model (4.22) in the next section). In such a situation, the minimum contrast functional $U_T(\theta)$, given by (3.5), depends only on the parameter (sub)vector θ and produces an estimator for θ only. This can be considered as the first step of the estimation procedure. An estimation of the remaining parameters, represented by the (sub)vector η, can be found in Anh et al. [6]; there the case of the vector η of dimension 2 was treated, and the estimation was based on the use of the spectral densities of orders 2 and 3. The estimation of η in the general case (that is, when η is of dimension $m \geqslant 3$) can be done analogously to Anh et al. [6]. We mention that the spectral estimates $\hat{\sigma}_{2,T}(w) = \int I_T^2(\lambda) w(\lambda)\, d\lambda$ and $\hat{\sigma}_{k,T}^{(1)}$, $\hat{\sigma}_{k,T}^{(2)}$, $k = 3, \ldots, m+1$, given by the formulae (3.10) and (3.11) need to be brought into the analysis in order to obtain the estimate for the vector η.

Remark 5. The proofs of our main results rely on large sample properties of the sample spectral functionals $J_k^T(\varphi)$ stated in Section 2 (see (2.5)). Their proofs are based on the representation of the cumulants of these functionals in the form of singular integrals and evaluation of their asymptotic behavior. This technique was elaborated in Bentkus [11,12] and Bentkus and Rutkauskas [15]. However, in our approach we formulate more general conditions on the second- and higher-order spectra as well as on the weight function φ; these conditions are formulated in such a way that the spectral densities and weight function are treated simultaneously and this allows us to avoid the use of such restrictive conditions as boundedness or square integrability of spectral densities or summability of cumulants. The conditions for sample spectral functionals induce the corresponding set of conditions needed to state the main results on consistency and asymptotic normality of our estimators. For the case $k = 2$, these conditions are basically analogous to those imposed on the second-order spectral density and smoothing function in Heyde and Gay [34] for linear processes (see Condition A and Theorem 1 there), although our set of conditions has been deduced from a different technique. For linear models and the case $k > 2$, and assuming that the weights are properly factorized, our conditions, which are of the form of integrability of products of spectral densities and weights, can be simplified. Namely, for analysis of the integrals of products of spectral densities used in our conditions, the generalized Hölder machinery can be used to reformulate the conditions in the form of integrability of spectral densities themselves. Here the approach of Avram and Taqqu [7] based on a generalization of the Hölder–Young inequality, which has been recently known as the Hölder–Brascamp–Lieb–Barthe inequality, is applicable. This approach also works for some particular models where spectral densities are factorizable. Of course, our set of conditions is satisfied for the models with bounded spectral densities, where weights should be used for the continuous-time case. In Anh et al. [6], the performance of the proposed estimation method has been illustrated for the cases $k = 2, 3$ with the use of simulated data.

4. Applications

In this section we discuss several situations in which the use of higher-order information is needed for statistical inference. We will not present the solutions to all the situations considered, but rather show certain ways in which our results can be exploited. For a function $f(\lambda_1, \ldots, \lambda_n)$ of n variables, $n \in \mathbb{N}$, its symmetrized version is given by

$$\underset{\lambda_1,\ldots,\lambda_n}{sym} f(\lambda_1, \ldots, \lambda_n) = \frac{1}{n!} \sum_{(i_1,\ldots,i_n)\in\pi_n} f\left(\lambda_{i_1}, \ldots, \lambda_{i_n}\right), \tag{4.1}$$

where π_n denotes the set of permutations of the numbers $\{1, \ldots, n\}$.

Example 1. Doppler processing using the trispectrum.

Dwyer [24] and Anderson et al. [2] investigated the use of higher-order information from signals which have undergone Gaussian amplitude modulation within the duration of the waveform. Such modulation is responsible for Doppler spreading of skywave sonar signal and may well contribute to observed Doppler spreading of skywave radar signal in the course of their passage through the turbulent ionospheric plasma. This approach is based on the model

$$Y(t) = Z(t)\, X(t), \quad t \in \mathbb{R},$$

where $Y(t)$ is the received signal, $X(t)$ is the waveform and $Z(t)$ is the amplitude modulation process, which is independent of $X(t)$. Noting that

$$\begin{aligned} c^Y(\tau) &= cov\,(Y(t), Y(t+\tau)) = cov\,(Z(t), Z(t+\tau))\, cov\,(X(t), X(t+\tau)) \\ &= c^Z(\tau)\, c^X(\tau), \end{aligned}$$

and choosing

$$Y(t) = Z(t)\, e^{-i(\lambda_0 t + \phi)}$$

with ϕ uniformly distributed on $[0, 2\pi]$, we get

$$c^Y(\tau) = c^Z(\tau)\, e^{i\lambda_0 \tau}.$$

We now assume that $Z(t)$ is a Gaussian Ornstein–Uhlenbeck process which is governed by the stochastic differential equation (SDE)

$$dZ(t) = -\alpha Z(t)\, dt + \alpha\, dW(t), \quad t \in \mathbb{R}, \tag{4.2}$$

where $\alpha > 0$ and $W(t)$ is a one-dimensional Brownian motion or Wiener process such that $EB(t) = 0$, $EB^2(t) = |t|$. In this case

$$c^Z(\tau) = \frac{\alpha}{2} e^{-\alpha|\tau|}, \quad \tau \in \mathbb{R}, \tag{4.3}$$

and the second-order spectral density of $Y(t)$ is given by

$$f_2^Y(\lambda) = f_2^Y(\lambda, \lambda_0) = \frac{1}{2\pi} \frac{\alpha^2}{(\lambda - \lambda_0)^2 + \alpha^2}, \quad \lambda \in \mathbb{R}, \tag{4.4}$$

where the most important parameter of Doppler effect is $\lambda_0 \in \mathbb{R}$. However, as $\alpha \to \infty$, from (4.3) and (4.4) we obtain

$$f_2^Y(\lambda) \to \frac{1}{2\pi}, \quad c^Z(\tau) \to \delta(\tau). \tag{4.5}$$

This represents generalized Gaussian white noise. That is, the signal is lost in stationary Gaussian noise.

On the other hand, the trispectrum $f_4(\lambda_1, \lambda_2, \lambda_3) = f_4(\lambda_1, \lambda_2, \lambda_3, \lambda_0)$ has an explicit form (see [2]) and $\lim_{\alpha\to\infty} f^Y(\lambda_1, \lambda_2, \lambda_2, \lambda_0)$ crucially depends on the parameter λ_0, yielding that the signal survives in the fourth order after modulation by Gaussian white noise. In fact, let us assume that $0 \leqslant \tau_1 \leqslant \tau_2 \leqslant \tau_3$. Then

$$
\begin{aligned}
&EY(t)\,Y(t+\tau_1)\,\overline{Y(t+\tau_2)Y(t+\tau_3)}\\
&\quad = E(Z(t)\,Z(t+\tau_1)\,Z(t+\tau_2)\,Z(t+\tau_3)\\
&\qquad \times e^{-i\left(\lambda_0 t+\phi\right)} e^{-i\left(\lambda_0(t+\tau_1)+\phi\right)} e^{i\left(\lambda_0(t+\tau_2)+\phi\right)} e^{i\left(\lambda_0(t+\tau_3)+\phi\right)})\\
&\quad = e^{-i\lambda_0\tau_1} e^{i\lambda_0\tau_2} e^{i\lambda_0\tau_3}[cov\,(Z(t)\,,Z(t+\tau_1))\,cov\,(Z(t+\tau_2)\,,Z(t+\tau_3))\\
&\qquad +cov\,(Z(t)\,,Z(t+\tau_2))\,cov\,(Z(t+\tau_1)\,,Z(t+\tau_3))\\
&\qquad +cov\,(Z(t)\,,Z(t+\tau_3))\,cov\,(Z(t+\tau_1)\,,Z(t+\tau_2))]\\
&\quad = e^{-i\lambda_0\tau_1} e^{i\lambda_0\tau_2} e^{i\lambda_0\tau_3}\\
&\qquad \times \left[\frac{\alpha}{2}\,e^{-\alpha\tau_1}\frac{\alpha}{2}\,e^{-\alpha(\tau_3-\tau_2)} + \frac{\alpha}{2}\,e^{-\alpha\tau_2}\frac{\alpha}{2}\,e^{-\alpha(\tau_3-\tau_1)} + \frac{\alpha}{2}\,e^{-\alpha\tau_3}\frac{\alpha}{2}\,e^{-\alpha(\tau_2-\tau_1)}\right],
\end{aligned}
$$

and

$$
\begin{aligned}
&cum\,(Y(t)\,,Y(t+\tau_1)\,,Y(t+\tau_2)\,,Y(t+\tau_3))\\
&\quad = e^{-i\lambda_0\tau_1}\frac{\alpha}{2}\,e^{-\alpha\tau_1}e^{i\lambda_0(\tau_3-\tau_2)}\frac{\alpha}{2}\,e^{-\alpha(\tau_3-\tau_2)}e^{2i\lambda_0\tau_2}\\
&\qquad + e^{i\lambda_0\tau_2}\frac{\alpha}{2}\,e^{-\alpha\tau_2}e^{i\lambda_0(\tau_3-\tau_1)}\frac{\alpha}{2}\,e^{-\alpha(\tau_3-\tau_1)} + e^{i\lambda_0\tau_3}\frac{\alpha}{2}\,e^{-\alpha\tau_3}e^{i\lambda_0(\tau_2-\tau_1)}\frac{\alpha}{2}\,e^{-\alpha(\tau_2-\tau_1)}\\
&\qquad - \left[\frac{\alpha}{2}\,e^{-\alpha\tau_1}e^{i\lambda_0\tau_1}\frac{\alpha}{2}\,e^{-\alpha(\tau_3-\tau_2)}e^{i\lambda_0(\tau_3-\tau_2)} + \frac{\alpha}{2}\,e^{-\alpha\tau_2}e^{i\lambda_0\tau_2}\frac{\alpha}{2}\,e^{-\alpha(\tau_3-\tau_1)}e^{i\lambda_0(\tau_3-\tau_1)}\right.\\
&\qquad \left. + \frac{\alpha}{2}\,e^{-\alpha\tau_3}e^{i\lambda_0\tau_3}\frac{\alpha}{2}\,e^{-\alpha(\tau_2-\tau_1)}e^{i\lambda_0(\tau_2-\tau_1)}\right].
\end{aligned}
\tag{4.6}
$$

The first term is equal to

$$
\begin{aligned}
&\frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} e^{i\lambda_1\tau_1}\frac{\alpha^2}{(\lambda_1+\lambda_0)^2+\alpha^2}\,e^{i\lambda_3(\tau_3-\tau_2)}\frac{\alpha^2}{(\lambda_3-\lambda_0)^2+\alpha^2}\,e^{i\lambda_2\tau_2}\\
&\quad \times\delta\,(\lambda_2-2\lambda_0)\;d\lambda_1\,d\lambda_2\,d\lambda_3\\
&\quad = \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} e^{i\lambda_1\tau_1+i\lambda_2\tau_2+i\lambda_3\tau_3}e^{-i\lambda_3\tau_2}\frac{\alpha^2}{(\lambda_1+\lambda_0)^2+\alpha^2}\,\frac{\alpha^2}{(\lambda_3-\lambda_0)^2+\alpha^2}\\
&\quad \times\delta\,(\lambda_2-2\lambda_0)\;d\lambda_1\,d\lambda_2\,d\lambda_3.
\end{aligned}
$$

Using the change of variables $\lambda_1=\widetilde{\lambda}_1$, $\lambda_2-\lambda_3=\widetilde{\lambda}_2$, $\lambda_3=\widetilde{\lambda}_3$, we obtain the first term of the trispectrum as

$$
\frac{1}{(2\pi)^3}\,\frac{\alpha^2}{(\lambda_1+\lambda_0)^2+\alpha^2}\,\frac{\alpha^2}{(\lambda_3-\lambda_0)^2+\alpha^2}\,\delta\,(\lambda_2+\lambda_3-2\lambda_0)\,. \tag{4.7}
$$

As $\alpha\to\infty$, this term of the trispectrum has the limit

$$
\frac{1}{(2\pi)^3}\,\delta\,(\lambda_2+\lambda_3-2\lambda_0)\,,
$$

which depends on λ_0. Formally, we can confirm this effect for the complete expression of $f_4\,(\lambda_1,\lambda_2,\lambda_3,\lambda_0)$. In fact, for $0\leqslant\tau_1\leqslant\tau_2\leqslant\tau_3$, analogously to the derivation of the term (4.7), we obtain from (4.6) the following expression for the first part of the trispectrum:

$$
\begin{aligned}
f_4^{(1)}(\lambda_1,\lambda_2,\lambda_3,\lambda_0) = \frac{1}{(2\pi)^3}\Bigg\{ & \frac{\alpha^2}{(\lambda_1+\lambda_0)^2+\alpha^2}\,\frac{\alpha^2}{(\lambda_3-\lambda_0)^2+\alpha^2}\,\delta(\lambda_2+\lambda_3-2\lambda_0)\\
& + \frac{\alpha^2}{(\lambda_3-\lambda_0)^2+\alpha^2}\,\frac{\alpha^2}{(\lambda_2-\lambda_0)^2+\alpha^2}\,\delta(\lambda_1+\lambda_3)\\
& + \frac{\alpha^2}{(\lambda_3-\lambda_0)^2+\alpha^2}\,\frac{\alpha^2}{(\lambda_2-\lambda_0)^2+\alpha^2}\,\delta(\lambda_1+\lambda_2)\\
& - \Bigg[\frac{\alpha^2}{(\lambda_1-\lambda_0)^2+\alpha^2}\,\frac{\alpha^2}{(\lambda_3-\lambda_0)^2+\alpha^2}\,\delta(\lambda_2+\lambda_3)\\
& + \frac{\alpha^2}{(\lambda_2-\lambda_0)^2+\alpha^2}\,\frac{\alpha^2}{(\lambda_3-\lambda_0)^2+\alpha^2}\,\delta(\lambda_1+\lambda_3)\\
& + \frac{\alpha^2}{(\lambda_3-\lambda_0)^2+\alpha^2}\,\frac{\alpha^2}{(\lambda_2-\lambda_0)^2+\alpha^2}\,\delta(\lambda_2+\lambda_1)\Bigg]\Bigg\}.
\end{aligned}
$$

As $\alpha \to \infty$, the last expression has the limit

$$
\frac{1}{(2\pi)^3}\left[\delta(\lambda_2+\lambda_3-2\lambda_0)-\delta(\lambda_2+\lambda_3)\right].
$$

To continue this procedure, we have to consider the following three regions: $\tau_1 \leqslant 0 \leqslant \tau_2 \leqslant \tau_3$; $\tau_1 \leqslant \tau_2 \leqslant 0 \leqslant \tau_3$ and $\tau_1 \leqslant \tau_2 \leqslant \tau_3 \leqslant 0$, and to perform similar computations. Then the fourth-order spectral density is obtained by summing up all the resulting terms obtained by considering all permutations of the subscripts 1–3. The result is that the parameter λ_0 is preserved in the fourth-order information even in Gaussian noise. Thus, we may use our minimum contrast method with $k = 4$ for the estimation of λ_0. This procedure is more stable than a procedure based on the second-order information in view of the danger of loss of the signal in Gaussian noise.

Remark. Some further applications of bispectral analysis to radar signals were given in Joury [41] and Pezeshki and Chandran [49]. The latter paper was devoted to applications of higher-order spectra to magnetically buckled beam.

Example 2. Bispectrum of a diffusion process with linear generator.

We use some ideas and facts from Yamada and Watanabe [55], Ikeda and Watanabe [40] and Iglói and Terdik [39]. Let us consider the SDE

$$
dY(t) = (\mu - 2\alpha Y(t))\,dt + 2\sigma\sqrt{Y(t)}\,dW(t), \quad t \in \mathbb{R}, \tag{4.8}
$$

where we assume that $Y(0) \geqslant 0$, $\mu > 0$, $\alpha \in \mathbb{R}$, $\sigma > 0$ and $W(t)$ is a standard Wiener process with mean zero and variance $EW^2(t) = |t|$. Eq. (4.8) has a unique strong solution [40, Theorem IV.3.2]. We call this solution $Y(t)$ a diffusion process with linear generator (DLG), since both coefficients of its generator

$$
A = (\mu - 2\alpha x)\frac{\partial}{\partial x} + 2\sigma^2 x \frac{\partial^2}{\partial x^2}
$$

are linear functions. We denote

$$d = \frac{\mu}{\sigma^2}.$$

The process $Y(t)$ is also called a Bessel process or square-root diffusion process. This process has recently been used in financial modelling (see [22]). If d is a non-negative integer and $\alpha > 0$, the DLG process is the radial part of a d-dimensional Ornstein–Uhlenbeck process, that is,

$$Y(t) = \sum_{i=1}^{d} X_i^2(t), \tag{4.9}$$

where $X_i(t), i = 1, \ldots, d$ are independent copies of a Gaussian Ornstein–Uhlenbeck process $X(t)$ with zero mean and covariance function

$$EX(t)X(t+\tau) = \frac{\sigma^2}{2\alpha} e^{-\alpha|\tau|} = \int_{\mathbb{R}} \cos(\lambda\tau) f_2^X(\lambda)\, d\lambda, \tag{4.10}$$

and second-order spectral density

$$f_2^X(\lambda) = \frac{1}{2\pi} \frac{\sigma^2}{\lambda^2 + \alpha^2}, \quad \lambda \in \mathbb{R} \tag{4.11}$$

(see [39]). Note that $X(t)$ is the Itô solution of the SDE

$$dX(t) = -\alpha X(t)\, dt + \sigma\, dW(t), \quad t \in \mathbb{R}, \tag{4.12}$$

where $W(t)$ is a standard Wiener process. From (4.9) we obtain that the marginal distribution of the process $Y(t)$ is a chi-square distribution. Moreover, it can be shown that for $\alpha > 0$ and $t \to \infty$ there exists a stationary distribution of the process $Y(t)$ (see [39]). The stationary distribution of a DLG process is a gamma distribution $\Gamma\left(d/2, \sigma^2/\alpha\right)$, that is, the density function of the marginal distribution of a stationary DLG process is of the form

$$g(x) = \frac{x^{(d/2)-1} e^{-x\alpha/\sigma^2}}{\left(\sigma^2/\alpha\right)^{d/2} \Gamma(d/2)}, \quad x > 0. \tag{4.13}$$

It follows from Propositions 8 and 9 of Iglói and Terdik [39] that the second-order spectral density of a stationary DLG process with marginal density (4.13) is of the form

$$f_2^Y(\lambda) = \frac{1}{2\pi} d \left(\frac{\alpha}{\sigma^2}\right)^{-2} \frac{2\alpha}{\lambda^2 + 4\alpha^2}, \quad \lambda \in \mathbb{R}, \tag{4.14}$$

while the third-order spectral density (bispectrum) takes the form

$$f_3^Y(\lambda_1, \lambda_2) = \left(\frac{1}{2\pi}\right)^2 d \left(\frac{\alpha}{\sigma^2}\right)^{-3} 3! \operatorname*{sym}_{\lambda_1,\lambda_2,\lambda_3} \left\{ \frac{1}{-i\lambda_1 + 2\alpha} \frac{1}{i\lambda_2 + 2\alpha} \right\},$$
$$(\lambda_1, \lambda_2) \in \mathbb{R}^2, \tag{4.15}$$

where $\lambda_1 + \lambda_2 + \lambda_3 = 0$ and the notation $sym_{\lambda_1,\lambda_2,\lambda_3}$ is defined by (4.1). This bispectrum can also be written as

$$f_3^Y(\lambda_1, \lambda_2) = \left(\frac{1}{2\pi}\right)^2 48d \left(\frac{\alpha}{\sigma^2}\right)^{-3} \frac{\alpha^2 \left(\lambda_1\lambda_2 + \lambda_1^2 + \lambda_2^2 + 12\alpha^2\right)}{\left(\lambda_1^2 + 4\alpha^2\right)\left(\lambda_2^2 + 4\alpha^2\right)\left((\lambda_1 + \lambda_2)^2 + 4\alpha^2\right)}. \tag{4.16}$$

Hence $f_3^Y(\lambda_1, \lambda_2)$ is a real-valued and positive function, which is quite unusual in the theory of higher-order spectral densities. Note that the second-order spectrum (4.14) of the stationary DLG process $Y(t)$ is the same as that of a Gaussian Ornstein–Uhlenbeck process $X(t)$ (see (4.11)) up to constants. Thus, we cannot use second-order information for testing the hypothesis: $\mathrm{H}_0 : Z(t) = X(t)$ against the alternative $\mathrm{H}_1 : Z(t) = Y(t)$, where $Z(t), 0 \leqslant t \leqslant T$ is an observation process. However, the third-order spectral densities of these two different processes $X(t)$ and $Y(t)$ are quite different. In fact $f_3^X(\lambda_1, \lambda_2) \equiv 0$ while $f_3^Y(\lambda_1, \lambda_2)$ is given by (4.16). Thus, we may apply our Theorems 1 and 2 with $k = 2$ and 3 to estimate the unknown parameter $\alpha > 0$. Then the width of the asymptotic confidence interval based on Theorem 2 permits us to construct a reasonable asymptotic test for checking the hypothesis H_0 against the alternative H_1.

Example 3. Bilinear stochastic systems with Brownian motion input.

Let us consider the SDE

$$dY(t) = (\mu + \alpha Y(t))\, dt + (\sigma + \beta Y(t))\, dW(t), \quad t \in \mathbb{R}, \tag{4.17}$$

where $W(t)$ is a standard Brownian motion and $\mu, \alpha, \sigma, \beta$ are constants. The process $Y(t)$ is a possible model for the interest rate (see [21]).

If $\mu = 0, \alpha < 0, \beta = 0$, the process $Y(t)$ is the Gaussian Ornstein–Uhlenbeck process with covariance function (4.10) and spectral density (4.11) in which α should be replaced by $-\alpha$.

We now consider the case where μ, α, σ and β are complex numbers (see [38]). The solution of (4.17) is known [42, p. 111]:

$$\begin{aligned} Y(t) = {} & e^{\left(\alpha-\beta^2/2\right)t+\beta W(t)}\Big[Y(0) + (\mu - \sigma\beta)\int_0^t e^{-\left(\alpha-\beta^2/2\right)s-\beta W(s)}\, ds \\ & +\sigma \int_0^t e^{-\left(\alpha-\beta^2/2\right)s-\beta W(s)}\, dW(s)\,. \end{aligned} \tag{4.18}$$

When $\beta = 0$, $\mathrm{Re}\,\alpha < 0$, $\sigma \neq 0$, this provides the stationary Gaussian Ornstein–Unlenbeck process (let us denote it by $Y_{\mathrm{G}}(t)$) with mean

$$EY_{\mathrm{G}}(t) = -\mu/\alpha,$$

covariance function

$$R(t) = \frac{|\sigma|^2}{2\,\mathrm{Re}\,\alpha} e^{\alpha t}, \quad t > 0,$$

$$R(t) = \overline{R(-t)}, \quad t < 0,$$

and second-order spectral density

$$f_2(\lambda) = -\frac{R(0)\,\mathrm{Re}\,\alpha}{\pi\,|i\lambda - \alpha|^2}, \quad \lambda \in \mathbb{R}. \tag{4.19}$$

When $|\mu|^2 + |\sigma|^2 > 0$, $\mu \neq 0$, $\beta \neq 0$, $\mathrm{Re}\,\alpha < 0$, $2\,\mathrm{Re}\,\alpha + |\beta|^2 < 0$, then there exists the stationary non-Gaussian solution of (4.17).

Assume that $\sigma = 0$ and $\mu \neq 0$. Then we obtain a stationary non-Gaussian process $Y_{\mathrm{NG}}(t)$ with

$$EY_{\mathrm{NG}}(t) = -\mu/\alpha,$$

covariance function

$$R(t) = \frac{-|\mu|^2 |\beta|^2}{|\alpha|^2 \left(2 \operatorname{Re} \alpha + |\beta|^2\right)} e^{\alpha t}, \quad t > 0,$$

$$R(t) = \overline{R(-t)}, \quad t < 0,$$

and second-order spectral density

$$f_2(\lambda) = -\frac{R(0) \operatorname{Re} \alpha}{\pi |i\lambda - \alpha|^2}, \quad \lambda \in \mathbb{R}. \tag{4.20}$$

We see that there is no difference between the second-order spectral density (4.19) of a Gaussian stationary Ornstein–Uhlenbeck process $Y_{\mathrm{G}}(t)$ and the second-order spectral density (4.20) of a non-Gaussian stationary process $Y_{\mathrm{NG}}(t)$ defined by (4.17) with $\sigma = 0$ (these spectral densities differ only in a multiplicative constant term represented by $R(0)$). Therefore, it is necessary to consider higher-order spectra for a testing of the hypothesis $\mathrm{H}_0 : Z(t) = Y_{\mathrm{NG}}(t)$ non-Gaussian, against the alternative $\mathrm{H}_1 : Z(t) = Y_{\mathrm{G}}(t)$, a Gaussian Ornstein–Uhlenbeck process. The higher-order spectral densities can be found from the chaotic representation of the process (4.18) [38]. Note that similar effects hold for fractional Brownian motion input [38]. In order to distinguish between the Gaussian (hypothesis H_1) and non-Gaussian (hypothesis H_0) cases, we can use such parameters as skewness

$$\gamma_3 = EZ^3(t) = c_3^Z(0,0) = \int_{\Lambda^2} f_2^Z(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2$$

and kurtosis

$$\gamma_4 = EZ^4(t) - 3EZ^2(t) = c_4^Z(0,0,0) = \int_{\Lambda^3} f_3^Z(\lambda_1, \lambda_2, \lambda_3)\, d\lambda_1\, d\lambda_2\, d\lambda_3.$$

These parameters are equal to zero under H_1. Consistent estimates of these parameters can be derived from our statistics (3.10) and (3.11) and their asymptotics (the results for the asymptotic distributions of the statistics (3.10), (3.11) for the case $k = 3$ can be extracted from Remark 5 of Anh et al. [6], and the extension to the case $k = 4$ is straightforward). For practical implementation, the model can be discretized, with the corresponding discretized versions of spectral densities constructed from well-known techniques; we can choose the weights $\equiv 1$. A complete solution of this kind of hypothesis testing problems needs further theoretical development, which was not a purpose of this paper.

Example 4. Non-Gaussian linear processes.

We consider a stationary non-Gaussian linear process

$$X(t) = m + \int_{\mathbb{R}} G_\theta(t-u)\, dL(u), \tag{4.21}$$

where m is a constant, $G_\theta(t)$ is a non-random memory function such that

$$\int_{\mathbb{R}} G_\theta^2(t)\, dt < \infty,$$

and $L(t)$ is a Lévy process with jth cumulant K_j (assumed to exist). Then [16, Example 3] the kth order spectral density of $X_\theta(t)$ is given by

$$f_k(\lambda_1, \ldots, \lambda_{k-1}, \theta) = (2\pi)^{-k+1} K_k g_\theta(i\lambda_1) \times g_\theta(i\lambda_2) \ldots g_\theta(i\lambda_{k-1}) g_\theta(-i(\lambda_1 + \cdots + \lambda_{k-1})), \quad k \geqslant 2, \tag{4.22}$$

where the Laplace transform

$$g_\theta(p) = \int_0^\infty e^{-pt} G_\theta(t)\, dt$$

of the memory function $G_\theta(t)$ is assumed to exist. Barndorff-Nielsen and Shephard [9,10] considered non-Gaussian Ornstein–Uhlenbeck-based models and their uses in financial econometrics. In these models, the memory function

$$G_\theta(t) = e^{-\theta t} \mathbf{1}_{(0,\infty)}(t), \quad \theta > 0.$$

They used a non-Gaussian Lévy process with positive jumps to model stochastic volatility. Further examples can be found in Anh et al. [3], who considered the Green function solutions of Langevin or fractional Langevin equations driven by Lévy noise $\dot{L}(t)$. The simplest case is the one-term equation

$$A \frac{d}{dt} X(t) + CX(t) = \dot{L}(t),$$

where A and C being constants. The Green function solution then takes the form (4.21) with the Green function

$$G(t) = \frac{1}{A} e^{-\frac{C}{A}t} \mathbf{1}_{(0,\infty)}(t).$$

Therefore, the spectral densities of the process $X(t)$ are of the form (4.22) with

$$g(p) = \frac{1}{Ap + C}.$$

In particular,

$$f_2(\lambda) = -\frac{\Psi^{(2)}(0)}{2\pi} \cdot \frac{1}{A^2\lambda^2 + C^2},$$

$$f_3(\lambda_1, \lambda_2) = \frac{\Psi^{(3)}(0)}{(2\pi)^2 i^3} \cdot \frac{1}{(Ai\lambda_1 + C)(Ai\lambda_2 + C)(-Ai(\lambda_1 + \lambda_2) + C)},$$

where $\Psi(\zeta) = \log E \exp\{i\zeta L(1)\}$, the characteristic exponent for the underlying Lévy process. Consider the case $A = 1$ and $\theta = C$:

$$f_2(\lambda) = -\frac{\Psi^{(2)}(0)}{2\pi} \cdot \frac{1}{\lambda^2 + C^2},$$

$$f_3(\lambda_1, \lambda_2) = \frac{\Psi^{(3)}(0)}{(2\pi)^2 i^3} \cdot \frac{(C - i\lambda_1)(C - i\lambda_2)(C + i(\lambda_1 + \lambda_2))}{(\lambda_1^2 + C^2)(\lambda_2^2 + C^2)((\lambda_1 + \lambda_2)^2 + C^2)}$$

$$= \frac{\Psi^{(3)}(0)}{(2\pi)^2 i^3} \cdot \frac{C(C^2+\lambda_1^2+\lambda_2^2+\lambda_1\lambda_2) - i\lambda_1\lambda_2(\lambda_1+\lambda_2)}{(\lambda_1^2+C^2)(\lambda_2^2+C^2)((\lambda_1+\lambda_2)^2+C^2)},$$

$$\vdots$$

$$f_q(\lambda_1,\ldots,\lambda_{q-1}) = \frac{\Psi^{(q)}(0)}{(2\pi)^{q-1} i^q} \cdot \frac{(C-i\lambda_1)\ldots(C-i\lambda_{q-1})(C+i(\lambda_1+\cdots+\lambda_{q-1}))}{(\lambda_1^2+C^2)\ldots(\lambda_{q-1}^2+C^2)((\lambda_1+\cdots+\lambda_{q-1})^2+C^2)}.$$

Denote

$$h_q(\eta) = \frac{\Psi^{(q)}(0)}{(2\pi)^{q-1} i^q},$$

where η is the parameter vector of the Lévy process. We may consider two situations:

(i) *Use of the second-order spectral density and the functional $U_T(\theta)$ with $k=2$.*

In this situation $\theta = C$, $U_T(\theta) = U_T(C)$. We can choose the weight function $w(\lambda) = \frac{1}{1+\lambda^2}$. The corresponding results provided in Section 3 are

1. The minimum contrast estimator

$$\widehat{C}_T = \arg\min U_T\ (C)$$

is a consistent estimator of the parameter C, and the estimator

$$J_{2,T}(w) = \int_{\mathbb{R}} I_2^T\ (\lambda)\, w\ (\lambda)\ d\lambda$$

is a consistent estimator for $J_2(w; C, \eta) = \int_{\mathbb{R}} f_2\ (\lambda, C, \eta)\, w\ (\lambda)\ d\lambda$.

2. Asymptotic normality:

$$T^{1/2}\left(\widehat{C}_T - C_0\right) \xrightarrow{d} N\left(0, a_0 s_0^{-2}\right),$$

where $a_0 = a(C_0, \eta_0)$, $s_0 = s(C_0, \eta_0)$ with

$$s(C,\eta) = \int_{\mathbb{R}} f_2\ (\lambda; C, \eta)\, w\ (\lambda)\ \frac{\partial^2}{\partial C^2} \log\psi\ (\lambda; C)\ d\lambda$$

$$= h_2(\eta)\sigma^2\ (C) \int_{\mathbb{R}} w\ (\lambda)\, \psi\ (\lambda; C)\ \frac{\partial^2}{\partial C^2} \log\psi\ (\lambda; C)\ d\lambda,$$

$$a\ (C,\eta) = 2\pi\left[h_4(\eta)\left(\sigma^2\ (C)\right)^2 \int_{\mathbb{R}^2} w\ (\lambda)\, w\ (\mu)\ \frac{\partial}{\partial C}\psi\ (\lambda; C)\ \frac{\partial}{\partial C}\psi\ (\mu; C)\ d\lambda\, d\mu \right.$$

$$\left. +2\left(h_2(\eta)\sigma^2\ (\theta)\right)^2 \int_{\mathbb{R}} (w\ (\lambda))^2 \left(\frac{\partial}{\partial C}\psi\ (\lambda; C)\right)^2 d\lambda\right].$$

(ii) *Use of the third-order spectral density and the functional $U_T(\theta)$ with $k=3$.*

In this situation $\theta = C$, $U_T(\theta) = U_T(C)$. We can choose the weight function $w_3^{(1)}(\lambda_1,\lambda_2) = \frac{1}{1+\lambda_1^2+\lambda_2^2}$, and consider the functional

$$U_T\ (C) = -\int_{\mathbb{R}^2} \operatorname{Re} I_3^T\ (\lambda_1,\lambda_2)\, w_3^{(1)}\ (\lambda_1,\lambda_2) \log\psi_1\ (\lambda_1,\lambda_2; C)\ d\lambda_1\, d\lambda_2,$$

$$\operatorname{Re} f_3(\lambda_1, \lambda_2, C, \eta) = h_3(\eta) \cdot \operatorname{Re} \frac{(C - i\lambda_1)(C - i\lambda_2)(C + i(\lambda_1 + \lambda_2))}{(\lambda_1^2 + C^2)(\lambda_2^2 + C^2)((\lambda_1 + \lambda_2)^2 + C^2)},$$

$$\int_{\mathbb{R}^2} \operatorname{Re} f_3 (\lambda_1, \lambda_2, C, \eta)\, w_3^{(1)} (\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2 = h_3(\eta) \sigma_1^2(C),$$

$$\psi_1 (\lambda_1, \lambda_2; C) = \frac{\operatorname{Re} f_3 (\lambda_1, \lambda_2, C, \eta)\, w_3^{(1)} (\lambda_1, \lambda_2)}{h_3(\eta)\sigma_1^2(C)}$$
$$= \frac{(C - i\lambda_1)(C - i\lambda_2)(C + i(\lambda_1 + \lambda_2))}{(\lambda_1^2 + C^2)(\lambda_2^2 + C^2)((\lambda_1 + \lambda_2)^2 + C^2)} \cdot \frac{1}{\sigma_1^2(C)}.$$

The corresponding results are

1. The minimum contrast estimator

$$\widehat{C}_T = \arg\min V_T (C)$$

is a consistent estimator of the parameter C, and the estimator

$$J_{3,T}(w) = \int_{\mathbb{R}^2} \operatorname{Re} \mathrm{I}_3^T (\lambda_1, \lambda_2)\, w_3^{(1)} (\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2$$

is a consistent estimator for $J_3(w; C, \eta) = \int_{\mathbb{R}^2} \operatorname{Re} f_3 (\lambda_1, \lambda_2)\, w_3^{(1)} (\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2 = h_3(\eta)\sigma_1^2(C)$.

2. Asymptotic normality:

$$T^{1/2} \left(\widehat{C}_T - C_0\right) \xrightarrow{d} N\left(0, a_0 s_0^{-2}\right),$$

where $a_0 = a(C_0, \eta_0)$, $s_0 = s(C_0, \eta_0)$ with

$$s(C, \eta) = h_3(\eta)\sigma^2 (C) \int_{\mathbb{R}^2} \left(\frac{\partial^2}{\partial C^2} \psi_1 (\lambda_1, \lambda_2; C) \right.$$
$$\left. - \frac{1}{\psi_1 (\lambda_1, \lambda_2; C)} \left(\frac{\partial}{\partial C} \psi_1 (\lambda_1, \lambda_2; C) \right)^2 \right) d\lambda_1\, d\lambda_2,$$

$$a (C, \eta) = \frac{h_6(\eta)}{2\,(h_3(\eta))^2} \left(\sigma^2 (C)\right)^2 \int_{\mathbb{R}^4} \left(1 - \frac{\operatorname{Im} f_3 (\lambda_1, \lambda_2) \operatorname{Im} f_3 (\lambda_4, \lambda_5)}{\operatorname{Re} f_3 (\lambda_1, \lambda_2) \operatorname{Re} f_3 (\lambda_4, \lambda_5)} \right)$$
$$\times \frac{\partial}{\partial C} \psi_1 (\lambda_1, \lambda_2; C) \frac{\partial}{\partial C} \psi_1 (\lambda_4, \lambda_5; C)\, d\lambda_1\, d\lambda_2\, d\lambda_4\, d\lambda_5$$
$$+ \frac{3h_2(\eta)h_4(\eta)}{2\,(h_3(\eta))^2} \left(\sigma^2 (C)\right)^2 \int_{\mathbb{R}^3} \left(1 - \frac{\operatorname{Im} f_3 (\lambda_1, \lambda_2) \operatorname{Im} f_3 (\lambda_1, -\mu)}{\operatorname{Re} f_3 (\lambda_1, \lambda_2) \operatorname{Re} f_3 (\lambda_1, -\mu)} \right)$$
$$\times \frac{\partial}{\partial C} \psi_1 (\lambda_1, \lambda_2; C) \frac{\partial}{\partial C} \psi_1 (\lambda_1, -\mu; C)\, d\lambda_1\, d\lambda_2\, d\mu$$
$$+ \frac{3h_2(\eta)h_4(\eta)}{(h_3(\eta))^2} \left(\sigma^2 (C)\right)^2 \int_{\mathbb{R}} \frac{\partial}{\partial C} \psi_1 (\lambda, 0; C)\, d\lambda \int_{\mathbb{R}^2} \frac{\partial}{\partial C} \psi_1 (\lambda_1, \lambda_2; C)\, d\lambda_1\, d\lambda_2$$
$$+ \frac{3\,(h_2(\eta))^3}{(h_3(\eta))^2} \left(\sigma^2 (C)\right)^2 \int_{\mathbb{R}^3} \left(1 + \left(\frac{\operatorname{Im} f_3 (\lambda_1, \lambda_2)}{\operatorname{Re} f_3 (\lambda_1, \lambda_2)} \right)^2 \right)$$

$$\times\left(\frac{\partial}{\partial C}\psi_1(\lambda_1,\lambda_2;C)\right)^2 d\lambda_1\,d\lambda_2$$
$$+\frac{3}{2}\left(\sigma^2(C)\right)^2\int_{\mathbb{R}^3}\left(1-\frac{\operatorname{Im} f_3(\lambda_1,\lambda_2)\operatorname{Im} f_3(\lambda_1+\lambda_2,-\mu)}{\operatorname{Re} f_3(\lambda_1,\lambda_2)\operatorname{Re} f_3(\lambda_1+\lambda_2,-\mu)}\right)$$
$$\times\frac{\partial}{\partial C}\psi_1(\lambda_1,\lambda_2;C)\,\frac{\partial}{\partial C}\psi_1(\lambda_1+\lambda_2,-\mu;C)\,d\lambda_1\,d\lambda_2\,d\mu$$
$$+\frac{9\,(h_2(\eta))^3}{2}\left(\sigma^2(C)\right)^2\left(\int_{\mathbb{R}}\frac{\partial}{\partial C}\psi_1(0,\lambda;C)\,d\lambda\right)^2.$$

Note that, in this particular case, the general formula (3.15) appears in a simple form with only the expression for the third-order spectral density $f_3(\lambda_1,\lambda_2)$ involved.

Example 5. Estimation of parameters of a non-Gaussian signal in the presence of Gaussian noise.

Let us consider the problem of estimation of an unknown parameter vector $\theta\in\Theta\subset\mathbb{R}^m$ of a strictly stationary non-Gaussian process $X_\theta(t)$ in the presence of a stationary Gaussian noise process $N_\zeta(t)$ which depends on a parameter vector $\zeta\in\Xi\subset\mathbb{R}^s$. In many situations θ and ζ have different natures. We also assume that the processes $X_\theta(t)$ and $N_\zeta(t)$ are independent.

Having observed the process

$$Y(t)=X_\theta(t)+N_\zeta(t) \tag{4.23}$$

over the interval $[0,T]$, we are interested in the statistical inference of the process $X_\theta(t)$. We assume that the spectral densities $f_k^X(\lambda_1,\ldots,\lambda_{k-1},\theta)$, $k\geqslant 2$ of the process $X_\theta(t)$ exist. The spectral densities $f_k^N(\lambda_1,\ldots,\lambda_{k-1},\zeta)$ of the Gaussian process $N_\zeta(t)$ exist and $f_k^N\equiv 0$ for $k\geqslant 3$. Therefore, the spectral densities of the process $Y(t)$ are of the form

$$f_2^Y(\lambda,\theta,\zeta)=f_2^X(\lambda,\theta)+f_2^N(\lambda,\zeta), \tag{4.24}$$

$$f_k^Y(\lambda_1,\ldots,\lambda_{k-1},\theta)=f_k^X(\lambda_1,\ldots,\lambda_{k-1},\theta),\quad k\geqslant 3. \tag{4.25}$$

We observe that the spectral densities f_k^Y, $k\geqslant 3$ do not depend on the parameter ζ of the noise process, which may have a rather complex structure. If we are interested in the estimation of the parameter vector θ of the signal, it is reasonable to use the analytical information on θ contained in the higher-order spectral densities $f_k^Y(\lambda_1,\ldots,\lambda_{k-1},\theta)$. We can apply Theorem 1 to construct a consistent estimator $\widehat{\theta}_T$ for the parameter θ (we may choose, for example, $k=3$). As for the asymptotic normality of the estimator $\widehat{\theta}_T$, we can apply Theorem 2. We should note, however, that the asymptotic variance in Theorem 2 depends on the second-order information through the expressions (3.15) and (2.7). Here we can proceed as follows. An option is to construct non-parametric estimates of the spectral functionals appearing in the expression for the covariance matrix, and therefore an estimate for this matrix is obtained. In this option, a further development of the theory is needed to extend the results of Section 2 to other types of spectral functionals, which will actually be linear functionals of spectral densities. On the other hand, having estimated θ, we can return to the estimation of the parameter vector ζ of the Gaussian process $N_\zeta(t)$ from the second-order spectral density (4.25) by substituting there $\widehat{\theta}_T$ instead of θ. We can apply again our minimum contrast functional, now with $k=2$, and obtain the estimate $\widehat{\zeta}_T$ of the parameter ζ. The question on the joint asymptotic normality of $\widehat{\theta}_T$ and $\widehat{\zeta}_T$ can be solved with the use of the delta-method (analogous to the situation considered in [6, Remark 5]).

We also observe that some functionals of the signal can be consistently estimated by using higher-order information. Suppose that in the model (4.23) the signal $X_\theta(t)$ is of the form (4.21), where the memory function is

$$G(t) = e^{-Ct}\mathbf{1}_{(0,\infty)}(t), \quad C > 0.$$

The third-order spectral density $f_3(\lambda_1, \lambda_2, c)$ of the model (4.23) does not depend on the parameter ζ of the noise process, while the second-order spectral density does (see (4.24)). By Theorems 1 and 2 the functional

$$J_{3,T}(w) = \int_{\mathbb{R}^2} \operatorname{Re} I_3^T(\lambda_1, \lambda_2)\, w_3^{(1)}(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2$$

is a consistent estimator for

$$J_3(w; c, \eta) = \int_{\mathbb{R}^2} \operatorname{Re} f_3(\lambda_1, \lambda_2, c)\, w_3^{(1)}(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2,$$

which does not depend on the parameter of the noise process. The weight function $w_3^{(1)}(\lambda_1, \lambda_2)$ is a function satisfying the conditions of Theorems 1 and 2.

A similar situation appears when we observe the process (4.23) in which the signal $X_\theta(t)$ has a bilinear structure:

$$X_\theta(t) = \int_{\mathbb{R}} a_\theta(t-u)\, dW(s) + \int_{\mathbb{R}^2} b_\theta(t-u, t-w)\, dW(u)\, dW(w), \tag{4.26}$$

where $W(t)$ is a standard Wiener process, $a_\theta(u)$ and $b_\theta(u, w)$ have Fourier transforms $A_\theta(\lambda)$, $B_\theta(\lambda_1, \lambda_2)$, respectively, and $b_\theta(u, w)$ is assumed to be symmetric in u and w for convenience. We also assume that

$$\int_{\mathbb{R}} a_\theta^2(u)\, du < \infty, \quad \int_{\mathbb{R}^2} b_\theta^2(u, w)\, du\, dw < \infty.$$

In the presence of a complicated Gaussian noise process $N_\zeta(t)$, the second-order spectral density $f_2^Y(\lambda)$ of the process (4.23) takes the form (4.24) which can be complex for many spectral densities $f_2^N(\lambda, \zeta)$. Also it depends on two parameters θ and ζ, while the third-order spectral density (bispectrum) $f_3^Y(\lambda_1, \lambda_2, \theta) = f_3^X(\lambda_1, \lambda_2, \theta)$ does not depend on the parameter ζ of the noise process $N_\zeta(t)$. This makes it quite reasonable to use the bispectrum for the estimation of the unknown parameter θ of a signal $X_\theta(t)$ of the form (4.26). Moreover, the bispectrum takes the analytical form

$$\begin{aligned} f_3(\lambda_1, \lambda_2, \theta) = \frac{2}{(2\pi)^3} [&A_\theta(\lambda_1)\, A_\theta(\lambda_2)\, B_\theta(-\lambda_1, -\lambda_2) \\ &+ A_\theta(\lambda_2)\, A_\theta(\lambda_3)\, B_\theta(-\lambda_2, -\lambda_3) + A_\theta(\lambda_3)\, A_\theta(\lambda_1)\, B_\theta(-\lambda_3, -\lambda_1) \\ &+ 8 \operatorname*{sym}_{\lambda_1,\lambda_2,\lambda_3} \int_{\mathbb{R}} B_\theta(\lambda, \lambda_1 - \lambda)\, B_\theta(\lambda_2 + \lambda, -\lambda)\, B_\theta(\lambda - \lambda_1, -\lambda - \lambda_2)\, d\lambda], \end{aligned} \tag{4.27}$$

(see [16, Example 4]) where the notation $sym_{\lambda_1,\lambda_2,\lambda_3}$ is defined by (4.1). If we choose some specific forms for the Fourier transforms $A_\theta(\lambda)$ and $B_\theta(\lambda_1, \lambda_2)$ of the transfer functions $a_\theta(u)$

and $b_\theta(u, w)$ in (4.27), we will obtain an explicit expression for the bispectrum (4.27) which then can be used in the minimization of the contrast functional (3.5) with $k = 3$, $p = q = \frac{1}{2}$.

From Theorem 1, we obtain consistent estimates of the functionals

$$J_3^{(1)}(\theta) = \int_{\mathbb{R}^2} \operatorname{Re} f_3(\lambda_1, \lambda_2, \theta)\, w_3^{(1)}(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2,$$

$$J_3^{(2)}(\theta) = \int_{\mathbb{R}^2} \operatorname{Im} f_3(\lambda_1, \lambda_2, \theta)\, w_3^{(2)}(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2$$

with $f_3(\lambda_1, \lambda_2, \theta)$ being given by (4.27) which does not depend on the noise process.

We now discuss briefly an extension of the above signal plus noise model. Let us consider a practically motivated model in signal processing presented in Nikias and Petropulu [48, Section 2.4.4], where the observed process is the sum of three independent processes:

$$S(t) = X(t) + Y(t) + Z(t),$$

where $X(t)$ being an AR(p_1) Gaussian process, $Y(t)$ an MA(q) non-Gaussian process with non-zero variance, skewness and kurtosis, and $Z(t)$ an AR(p_2) non-Gaussian process with non-zero variance and kurtosis, but zero skewness. Then we have the second-order spectrum

$$f_2^S(\lambda) = f_2^X(\lambda) + f_2^Y(\lambda) + f_2^Z(\lambda),$$

and the third-order and fourth-order spectral densities are of the form

$$f_3^S(\lambda_1, \lambda_2) = f_3^Y(\lambda_1, \lambda_2) \quad \text{and} \quad f_4^S(\lambda_1, \lambda_2, \lambda_3) = f_4^Y(\lambda_1, \lambda_2, \lambda_3) + f_4^Z(\lambda_1, \lambda_2, \lambda_3),$$

respectively. Therefore we can estimate the parameters of the non-Gaussian process $Y(t)$ from the third-order spectral density, then the parameters of $Z(t)$ from the fourth-order spectrum. The remaining parameters, namely those of $X(t)$, can be estimated from the second-order spectral density.

5. Proofs

The proofs of the results of the present paper are based on the properties of the multidimensional kernels of Fejér type $\Phi_k^T(u_1, \ldots, u_{k-1})$ (see Appendix A), the formulae relating moments and cumulants and the formulae giving expressions for the cumulants of products of random variables via products of cumulants of the individual variables (see Appendix B). We present the proofs for continuous-time case. The proofs use the following formulae for the cumulants of the finite Fourier transform $d_T(\lambda)$, $\lambda \in \mathbb{R}$:

$$\begin{aligned}
&cum\left(d_T(\alpha_1), \ldots, d_T(\alpha_k)\right) \\
&\quad = \int_{[0,T]^k} e^{-i\Sigma_1^k \alpha_j t_j}\, cum\left(Y(t_1), \ldots, Y(t_k)\right) dt_1 \ldots dt_k \\
&\quad = \int_{\mathbb{R}^{k-1}} f_k\left(\gamma_1, \ldots, \gamma_{k-1}\right) \int_{[0,T]^k} \exp\{it_1\left(\gamma_1 - \alpha_1\right)\} \ldots \exp\{it_{k-1}\left(\gamma_{k-1} - \alpha_{k-1}\right)\} \\
&\qquad \times \exp\left\{it_k\left(-\sum_{j=1}^{k-1}\gamma_j - \alpha_k\right)\right\} dt_1 \ldots dt_k\, d\gamma_1 \ldots d\gamma_{k-1}
\end{aligned}$$

$$= \int_{\mathbb{R}^{k-1}} f_k\left(\gamma_1, \ldots, \gamma_{k-1}\right) \phi_1^T\left(\gamma_1 - \alpha_1\right) \ldots \phi_1^T\left(\gamma_{k-1} - \alpha_{k-1}\right)$$

$$\times \phi_1^T\left(-\sum_1^{k-1} \gamma_j - \alpha_k\right) d\gamma_1 \ldots d\gamma_{k-1}, \tag{5.1}$$

where

$$\phi_1^T(\lambda) = \int_0^T e^{it\lambda}\, dt = \frac{\sin(T\lambda/2)}{\lambda/2} e^{i(T/2)\lambda}. \tag{5.2}$$

Note that if $\sum_{j=1}^k \lambda_j = 0$, then

$$\frac{1}{(2\pi)^{k-1} T} \prod_{j=1}^k \phi_1^T\left(\lambda_j\right) = \Phi_k^T\left(\lambda_1, \ldots, \lambda_{k-1}\right), \tag{5.3}$$

where $\Phi_k^T(\lambda_1, \ldots, \lambda_{k-1})$ is the multidimensional kernel of Fejér type (see Appendix A). Therefore for the case when $\sum_{j=1}^k \alpha_j = 0$, the formula (5.1) can be written as

$$\frac{1}{(2\pi)^{k-1} T} cum\left(d_T(\alpha_1), \ldots, d_T(\alpha_k)\right)$$

$$= \int_{\mathbb{R}^{k-1}} f_k\left(\gamma_1, \ldots, \gamma_{k-1}\right) \Phi_k^T\left(\gamma_1 - \alpha_1, \ldots, \gamma_{k-1} - \alpha_{k-1}\right) d\gamma_1 \ldots d\gamma_{k-1}$$

$$= \int_{\mathbb{R}^{k-1}} \Phi_k^T\left(u_1, \ldots, u_{k-1}\right) f_k\left(u_1 + \alpha_1, \ldots, u_{k-1} + \alpha_{k-1}\right) du_1 \ldots du_{k-1}. \tag{5.4}$$

The proofs are given here for the continuous-time case. The discrete-time case can be treated in a similar manner with the kernels $\Phi_k^T(u_1, \ldots, u_{k-1})$ replaced by the kernels $\Psi_k^T(u_1, \ldots, u_{k-1})$ (see Appendix A).

Proof of Lemma 1. In view of the formulae (B.3) and (5.4) we have

$$E J_k^T(\varphi) = \int_{\mathbb{R}^{k-1}} \varphi(\lambda) \psi(\lambda) \frac{1}{(2\pi)^{k-1} T} E\left(\prod_{i=1}^k d_T(\lambda_i)\right) \delta\left(\sum_{i=1}^k \lambda_i\right) d\lambda'$$

$$= \frac{1}{(2\pi)^{k-1} T} \int_{\mathbb{R}^{k-1}} \varphi(\lambda) \psi(\lambda)$$

$$\times \sum_{\substack{v=(v_1,\ldots,v_p) \\ \text{partitions of } (1,\ldots,k)}} \int_{\mathbb{R}^{k-p}} f_{|v_1|}\left(\gamma_j, j \in \tilde{v}_1\right) \ldots f_{|v_p|}\left(\gamma_j, j \in \tilde{v}_p\right)$$

$$\times \prod_{j=1}^k \phi_1^T\left(\gamma_j - \lambda_j\right) \prod_{l=1}^p \delta\left(\sum_{j \in v_l} \gamma_j\right) d\gamma' \delta\left(\sum_{i=1}^k \lambda_i\right) d\lambda'$$

$$= \int_{\mathbb{R}^{k-1}} \Phi_k^T(u)\, G_k(u)\, du',$$

where

$$G_k(u) = G_k(u_1, \ldots, u_k; \varphi, \psi)$$
$$= \sum_{v=(v_1,\ldots,v_p)} \int_{\mathbb{R}^{k-p}} \prod_{l=1}^{p} f_{|v_l|}\left(\lambda_j + u_j, j \in \tilde{v}_l\right)$$
$$\times H_k(\lambda) \prod_{l=1}^{p-1} \delta\left(\sum_{j\in v_l}\left(\lambda_j + u_j\right)\right) \delta\left(\sum_1^k \lambda_i\right) d\lambda',$$
$$H_k(\lambda) = \varphi(\lambda)\,\psi(\lambda).$$

Due to the presence of the function $\psi(\lambda)$ we have

$$G_k(0) = \int_{\mathbb{R}^{k-1}} f_k(\lambda)\, H_k(\lambda)\, \delta\left(\sum_{i=1}^k \lambda_i\right) d\lambda'.$$

In view of the properties of the kernels $\Phi_k^T(u)$ (see Appendix A) we have that if the function $G_k(u)$ is bounded and continuous at $u = 0$, then

$$EJ_k^T(\varphi) \to \int_{\mathbb{R}^{k-1}} f_k(\lambda)\, H_k(\lambda)\, \delta\left(\sum_{i=1}^k \lambda_i\right) d\lambda'$$

as $T \to \infty$. $\square$

Proof of Lemma 2. Using the formulae (B.2), (B.4) and (5.4) we have

$$cov\left(J_k^T(\varphi_1), J_k^T(\varphi_2)\right)$$
$$= cum\left(J_k^T(\varphi_1), \overline{J_k^T(\varphi_2)}\right)$$
$$= \frac{1}{(2\pi)^{2(k-1)}T^2} \int_{\mathbb{R}^{2(k-1)}} \varphi_1(\lambda_1, \ldots, \lambda_k)\, \bar{\varphi}_2(-\lambda_{k+1}, \ldots, -\lambda_{2k})$$
$$\times \psi(\lambda_1, \ldots, \lambda_k)\, \bar{\psi}(-\lambda_{k+1}, \ldots, -\lambda_{2k})\, \delta\left(\sum_{i=1}^k \lambda_i\right) \delta\left(\sum_{i=k+1}^{2k} \lambda_i\right)$$
$$\times cum\left(d_T(\lambda_1)\ldots d_T(\lambda_k), d_T(\lambda_{k+1})\ldots d_T(\lambda_{2k})\right) d\lambda'$$
$$= \frac{1}{(2\pi)^{2(k-1)}T^2} \int_{\mathbb{R}^{2(k-1)}} H_{2k}(\lambda) \sum_{v=(v_1,\ldots,v_p)} \int_{\mathbb{R}^{2k-p}} \prod_{l=1}^{p} f_{|v_l|}\left(\gamma_j, j \in \tilde{v}_l\right)$$
$$\times \prod_{j=1}^{2k} \phi_1^T\left(\gamma_j - \lambda_j\right) \prod_{l=1}^{p} \delta\left(\sum_{j\in v_l} \gamma_j\right) d\gamma'\, \delta\left(\sum_{j=1}^k \lambda_j\right) \delta\left(\sum_{j=k+1}^{2k} \lambda_j\right) d\lambda'$$
$$= \frac{2\pi}{T} \int_{\mathbb{R}^{2k-1}} \Phi_{2k}^T(u)\, G_{2k}(u)\, du',$$

where

$$G_{2k}(u) = \sum_{\substack{\nu=(\nu_1,\ldots,\nu_p)\\ 1\leqslant p\leqslant k}} \int_{\mathbb{R}^{2k-p-1}} \prod_{l=1}^{p} f_{|\nu_l|}\left(u_j+\lambda_j,\, j\in\tilde{\nu}_l\right) H_{2k}(\lambda)$$

$$\times\delta\left(\sum_{j=1}^{k}\lambda_j\right)\delta\left(\sum_{j=k+1}^{2k}\lambda_j\right)\prod_{l=1}^{p-1}\delta\left(\sum_{j\in\nu_l}(\lambda_j+u_j)\right)d\lambda',$$

with

$$H_{2k}(\lambda) = \varphi_1(\lambda_1,\ldots,\lambda_k)\,\bar{\varphi}_2(-\lambda_{k+1},\ldots,-\lambda_{2k})\,\psi(\lambda_1,\ldots,\lambda_k)\,\bar{\psi}(-\lambda_{k+1},\ldots,-\lambda_{2k}),$$

and the sum is taken over all indecomposable partitions of the table

1	...	k
$k+1$	...	$2k$

(see Appendix B). The statement of the lemma now follows from Lemmas A.1 and A.2 (see Appendix A).

We omit the proofs of the results for higher order cumulants (Lemmas 4 and 5) as they are the same as those for Lemmas 1 and 2. □

Proof of Lemma 6. Let us fix arbitrary constants $c_1,\ldots,c_m$ and consider the variable

$$Y_T = \sum_{j=1}^{m} c_j\left[J_k^T\left(\varphi_j\right) - E J_k^T\left(\varphi_j\right)\right]$$
$$= \int_{\mathbb{R}^{k-1}} F(\lambda)\, I_k^T(\lambda)\,\psi(\lambda)\,d\lambda' - E\int_{\mathbb{R}^{k-1}} F(\lambda)\, I_k^T(\lambda)\,\psi(\lambda)\,d\lambda'$$

with $F(\lambda) = \sum_{j=1}^{m} c_j\varphi_j(\lambda)$. Using Lemmas 1–4, we can show that the random variable $T^{1/2}Y_T$ tends in distribution as $T\to\infty$ to a normal random variable with mean zero and variance s^2 given by

$$s^2 = 2\pi G_{2k}(0; F, F, \psi) = \sum_{i,j=1}^{m} c_j\bar{c}_j 2\pi G_{2k}(0;\varphi_i,\varphi_j,\psi) = \sum_{i,j=1}^{m} c_j\bar{c}_j w_{ij},$$

which implies (2.10).

The convergence (2.11) follows from (2.10) if assumption III holds true. □

Proof of Theorem 1. In view of assumptions V, VIII(i) and Lemma 3, we obtain

$$U_T(\theta)\to U(\theta) \quad \text{in } P_0\text{-probability}, \tag{5.5}$$

hence

$$U_T(\theta) - U_T(\theta_0) \to U(\theta) - U(\theta_0) = K(\theta_0;\theta).$$

By Jensen's inequality and the relations (3.2)–(3.4)

$$
\begin{aligned}
-K(\theta_0;\theta) &= p\int_{\mathbb{R}^{k-1}} f_k^{(1)}(\lambda;\theta_0)\log\frac{\psi_k^{(1)}(\lambda;\theta)}{\psi_k^{(1)}(\lambda;\theta_0)}w^{(1)}(\lambda)\,w_0(\lambda)\,d\lambda' \\
&\quad + q\int_{\mathbb{R}^{k-1}} f_k^{(2)}(\lambda;\theta_0)\log\frac{\psi_k^{(2)}(\lambda;\theta)}{\psi_k^{(2)}(\lambda;\theta_0)}w^{(2)}(\lambda)\,w_0(\lambda)\,d\lambda' \\
&= p\sigma_k^{(1)}(\theta_0)\int_{\mathbb{R}^{k-1}} \psi_k^{(1)}(\lambda;\theta_0)\log\frac{\psi_k^{(1)}(\lambda;\theta)}{\psi_k^{(1)}(\lambda;\theta_0)}w_0(\lambda)\,d\lambda' \\
&\quad + q\sigma_k^{(2)}(\theta_0)\int_{\mathbb{R}^{k-1}} \psi_k^{(2)}(\lambda;\theta_0)\log\frac{\psi_k^{(2)}(\lambda;\theta)}{\psi_k^{(2)}(\lambda;\theta_0)}w_0(\lambda)\,d\lambda' \\
&\leqslant p\sigma_k^{(1)}(\theta_0)\log\int_{\mathbb{R}^{k-1}} \psi_k^{(1)}(\lambda;\theta_0)\,w_0(\lambda)\,d\lambda' \\
&\quad + q\sigma_k^{(2)}(\theta_0)\log\int_{\mathbb{R}^{k-1}} \psi_k^{(2)}(\lambda;\theta_0)\,w_0(\lambda)\,d\lambda' = 0,
\end{aligned}
$$

that is, $K(\theta_0;\theta)\geqslant 0$ and $K(\theta_0;\theta)>0$ if

$$\psi_k^{(i)}(\lambda;\theta_0)\not\equiv\psi_k^{(i)}(\lambda;\theta)\quad \text{for all } \theta\neq\theta_0 \text{ a.e. w.r.t. Lebesgue measure } i=1,2.$$

Further, to prove the consistency of the estimator $\widehat{\theta}_T$ given by (3.9), we need to show that the convergence (5.5) holds uniformly in $\theta\in\Theta$. Here, we use the conditions IX and X.

If the condition IX holds, let $\eta(\varepsilon)=\min\left\{\eta_1(\varepsilon),\eta_2(\varepsilon)\right\}$, with $\eta_i(\varepsilon)$ being the modulus of continuity of the function $h_k^{(i)}(\lambda;\theta)$, $i=1,2$. Then, we have

$$
\begin{aligned}
&\sup\left\{\left|U_T(\theta_1)-U_T(\theta_2)\right|,\theta_1,\theta_2\in\Theta,|\theta_1-\theta_2|\leqslant\eta(\varepsilon)\right\} \\
&\leqslant\varepsilon\left[p\int_{\mathbb{R}^{k-1}}\operatorname{Re} I_k^T(\lambda)\frac{w^{(1)}(\lambda)}{v_1(\lambda)}w_0(\lambda)\,d\lambda'+q\int_{\mathbb{R}^{k-1}}\operatorname{Im} I_k^T(\lambda)\frac{w^{(2)}(\lambda)}{v_2(\lambda)}w_0(\lambda)\,d\lambda'\right].
\end{aligned}
\tag{5.6}
$$

Now, in view of IX(ii) and Lemma 3, we obtain that

$$\int_{\mathbb{R}^{k-1}} I_k^T(\lambda)\frac{w^{(i)}(\lambda)}{v_i(\lambda)}w_0(\lambda)\,d\lambda'\to\int_{\mathbb{R}^{k-1}} f_k(\lambda;\theta_0)\frac{w^{(i)}(\lambda)}{v_i(\lambda)}w_0(\lambda)\,d\lambda' \tag{5.7}$$

in probability; from (5.7), we can conclude that

$$\int_{\mathbb{R}^{k-1}} I_k^T(\lambda)\frac{w^{(i)}(\lambda)}{v_i(\lambda)}w_0(\lambda)\,d\lambda'=O_p(1),$$

which implies that the expression in the square brackets on the right-hand side of (5.6) is $O_p(1)$. This completes the proof. □

Proof of Theorem 2. From Taylor's formula, we have

$$\nabla_\theta U_T\left(\widehat{\theta}_T\right)=\nabla_\theta U_T(\theta_0)+\nabla_\theta\nabla_\theta' U_T\left(\theta_T^*\right)\left(\widehat{\theta}_T-\theta_0\right),$$

where $\left|\theta_T^* - \theta_0\right| < \left|\widehat{\theta}_T - \theta_0\right|$;

$$\nabla_\theta U_T(\theta) = -\left(p \int_{\mathbb{R}^{k-1}} \operatorname{Re} I_k^T(\lambda)\, w^{(1)}(\lambda)\, w_0(\lambda)\, \nabla_\theta \log \psi_k^{(1)}(\lambda;\theta)\, d\lambda' \right.$$
$$\left. + q \int_{\mathbb{R}^{k-1}} \operatorname{Im} I_k^T(\lambda)\, w^{(2)}(\lambda)\, w_0(\lambda)\, \nabla_\theta \log \psi_k^{(2)}(\lambda;\theta)\, d\lambda' \right),$$

$$\nabla_\theta \nabla'_\theta U_T(\theta) = -\left(p \int_{\mathbb{R}^{k-1}} \operatorname{Re} I_k^T(\lambda)\, w^{(1)}(\lambda)\, w_0(\lambda) \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log \psi_k^{(1)}(\lambda;\theta)\right)_{i,j=\overline{1,m}} d\lambda' \right.$$
$$\left. + q \int_{\mathbb{R}^{k-1}} \operatorname{Im} I_k^T(\lambda)\, w^{(2)}(\lambda)\, w_0(\lambda) \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log \psi_k^{(2)}(\lambda;\theta)\right)_{i,j=\overline{1,m}} d\lambda' \right).$$

For sufficiently large T, we have

$$\nabla_\theta U_T(\theta_0) = -\nabla_\theta \nabla'_\theta U_T\left(\theta_T^*\right)\left(\widehat{\theta}_T - \theta_0\right). \tag{5.8}$$

If we can show that as $T \to \infty$

(1)

$$\nabla_\theta \nabla'_\theta U_T\left(\theta_T^*\right) \to S_k(\theta_0) \tag{5.9}$$

in P_0-probability, where the matrix $S_k(\theta_0)$ is given by (3.14), and

(2)

$$T^{1/2} \nabla_\theta U_T(\theta_0) \xrightarrow{d} N_m(0, A_k(\theta_0)), \tag{5.10}$$

where the matrix $A_k(\theta_0)$ is given by (3.15), then by Slutsky's arguments, the relation (3.16) is a consequence of (5.8)–(5.10). More precisely, the following multidimensional version of Slutsky's lemma (see, for example, [51]) is of use here.

Slutsky's Lemma. *Let $\{\xi_T\}$ and $\{\eta_T\}$ be families of random vectors in R^m, and let $\xi_T \xrightarrow{d} \xi$, $\eta_T \xrightarrow{P} \eta = \left(\eta_1, \ldots, \eta_m\right)$ as $T \to \infty$, where $\eta_1, \ldots, \eta_m$ are constants. Let $\{\zeta_T\}$ be a family of random matrices which elements tends in probability to elements of non-singular matrix ζ as $T \to \infty$. Then as $T \to \infty$: $\xi_T + \eta_T \xrightarrow{d} \xi + \eta$, $\zeta_T \xi_T \xrightarrow{d} \zeta\xi$.*

Let us prove (5.9) and (5.10).

In view of the assumptions XI(i), (ii), we have by Lemma 3

$$\int_{\mathbb{R}^{k-1}} I_k^T(\lambda)\, w^{(l)}(\lambda)\, w_0(\lambda) \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log \psi_k^{(l)}(\lambda;\theta)\right) d\lambda'$$
$$\xrightarrow{P_0} \int_{\mathbb{R}^{k-1}} f_k(\lambda;\theta_0)\, w^{(l)}(\lambda)\, w_0(\lambda) \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log \psi_k^{(l)}(\lambda;\theta)\right) d\lambda',$$

where $i, j = 1, \ldots, m; l = 1, 2$, which implies (5.9). Further, we are interested in the limiting distribution of the vector $T^{1/2}\nabla_\theta U_T(\theta_0)$ which can be written in the form

$$
\begin{aligned}
T^{1/2}\nabla_\theta U_T(\theta_0) &= T^{1/2}\left(p \operatorname{Re} J_k^T\left(g_k^{(i)}\right) + q \operatorname{Im} J_k^T\left(g_k^{(m+i)}\right)\right)_{i=1,\ldots,m} \\
&= T^{1/2}(p(\operatorname{Re} J_k^T(g_k^{(i)}) - \operatorname{Re} J_k(g_k^{(i)})) \\
&\quad + q(\operatorname{Im} J_k^T(g_k^{(m+i)}) - \operatorname{Im} J_k(g_k^{(m+i)})))_{i=1,\ldots,m}, \qquad (5.11)
\end{aligned}
$$

where the last equality is due to (3.8). Under the conditions of the theorem, we have by Lemma 6 as $T \to \infty$

$$
T^{1/2}\left(\{J_k^T(g_k^{(i)})\} - \{J_k(g_k^{(i)})\}\right) \xrightarrow{d} \zeta,
$$

where

$$
\begin{aligned}
\{J_k^T(g_k^{(i)})\} &= \left(J_k^T\left(g_k^{(1)}\right), \ldots J_k^T\left(g_k^{(2m)}\right)\right)', \\
\{J_k(g_k^{(i)})\} &= \left(J_k\left(g_k^{(1)}\right), \ldots J_k\left(g_k^{(2m)}\right)\right)',
\end{aligned}
$$

and $\zeta = (\zeta_1, \ldots, \zeta_{2m})'$ is a complex-valued Gaussian vector with mean zero and second-order moments

$$
E\zeta_i\overline{\zeta}_j = 2\pi G_{2k}\left(0; g_k^{(i)}, g_k^{(j)}, w_0\right), \quad i, j = 1, \ldots, 2m,
$$

with $G_{2k}\left(u; \varphi_1, \varphi_2, \psi\right)$ being given by the formula (2.7) and the functions $g_k^{(i)}, i = 1, \ldots, 2m$, being given by (3.13). Therefore, we can conclude that the vector (5.11) converges in distribution to an m-dimensional Gaussian vector with zero mean and second-order moments equal to $a_{ij}^{(k)}(\theta_0)$. This completes the proof of the theorem. $\square$

Proofs of Theorems 3 and 4 are analogous to those of Theorems 1 and 2.

Acknowledgments

The authors are grateful to the referees for their detailed comments and suggestions to improve the paper.

Appendix A. Multidimensional kernels of Fejér type

In the proofs, we use a technique based on the results of Ibragimov [36], Bentkus [11,12], Bentkus and Rutkauskas [15] about multidimensional kernels of Fejér type $\Phi_k^T(u_1, \ldots, u_{k-1})$, $(u_1, \ldots, u_{k-1}) \in \Lambda^{k-1}$, expressed by the formula

$$
\begin{aligned}
\Phi_k^T(u_1, \ldots, u_{k-1}) &= \frac{1}{(2\pi)^{k-1}T}\frac{\sin(Tu_1/2)}{u_1/2}\cdots\frac{\sin(Tu_{k-1}/2)}{u_{k-1}/2}\frac{\sin(T(u_1+\cdots+u_{k-1})/2)}{(u_1+\cdots+u_{k-1})/2} \\
&= \frac{1}{(2\pi)^{k-1}T}\int_{[0,T]^k}\exp\left\{i\Sigma_{j=1}^{k-1}t_j u_j - it_k\Sigma_{j=1}^{k-1}u_j\right\} dt_1\ldots dt_k \qquad (A.1)
\end{aligned}
$$

in the continuous case $\Lambda^{k-1} = \mathbb{R}^{k-1}$, and by the formula

$$\Phi_k^T(u_1, \dots, u_{k-1}) = \frac{1}{(2\pi)^{k-1} T} \frac{\sin(Tu_1/2)}{\sin(u_1/2)} \cdots \frac{\sin(Tu_{k-1}/2)}{\sin(u_{k-1}/2)} \frac{\sin(T(u_1 + \dots + u_{k-1})/2)}{\sin((u_1 + \dots + u_{k-1})/2)}$$

$$= \frac{1}{(2\pi)^{k-1} T} \sum_{t_1, \dots, t_k = \overline{1,T}} \exp\left\{ i \Sigma_{j=1}^{k-1} t_j u_j - i t_k \Sigma_{j=1}^{k-1} u_j \right\} \tag{A.2}$$

for the discrete case $\Lambda^{k-1} = (-\pi, \pi]^{k-1}$. Lemma A.1 shows that these kernels have properties which make them similar to the Fejér kernels

$$\frac{1}{2\pi T} \frac{\sin^2(Tx/2)}{(x/2)^2}, \quad x \in \mathbb{R} \quad \text{and} \quad \frac{1}{2\pi T} \frac{\sin^2(Tx/2)}{\sin^2(x/2)}, \quad x \in (-\pi, \pi].$$

Lemma A.1. *The kernel* $\Phi_k^T(u_1, \dots, u_{k-1})$, *with* $T \in (0, \infty)$ *and* $k \geqslant 2$ *an arbitrary integer, has the following properties*:

(1)

$$\sup_T \int_{\Lambda^{k-1}} \left| \Phi_k^T(u_1, \dots, u_{k-1}) \right| du_1 \dots du_{k-1} < \infty;$$

(2)

$$\int_{\Lambda^{k-1}} \Phi_k^T(u_1, \dots, u_{k-1}) \, du_1 \dots du_{k-1} = 1;$$

(3) *for arbitrary* $\varepsilon > 0$

$$\lim_{T \to \infty} \int_{\Lambda^{k-1} \setminus \{|u| \leqslant \varepsilon\}} \left| \Phi_k^T(u_1, \dots, u_{k-1}) \right| du_1 \dots du_{k-1} = 0,$$

where $\{|u| \leqslant \varepsilon\} = \left\{ (u_1, \dots, u_{k-1}) : |u_j| \leqslant \varepsilon, \ j = 1, \dots, k-1 \right\}$.

As a consequence of Lemma A.1, we have the following.

Lemma A.2. *Let a function* $G(u_1, \dots, u_{k-1})$ *be bounded and continuous at the point* $(u_1, \dots, u_{k-1}) = (0, \dots, 0)$. *Then*

$$\lim_{T \to \infty} \int_{\Lambda^{k-1}} \Phi_k^T(u_1, \dots, u_{k-1}) \, G(u_1, \dots, u_{k-1}) \, du_1 \dots du_{k-1} = G(0, \dots, 0).$$

The kernels $\Phi_k^T(u_1, \dots, u_{k-1})$, $(u_1, \dots, u_{k-1}) \in \Lambda^{k-1}$, in continuous and discrete time settings were considered, for example, in Bentkus [12,11], and were shown to satisfy the above conditions 1–3 of Lemma A.1.

Appendix B. Cumulants

Following Leonov and Shiryaev [47] (see also [19,13]), we describe an algorithm for the calculation of the joint cumulants of polynomial functions of random variables. This algorithm has been used extensively in the proofs of the present paper.

Consider a (not necessarily rectangular) two-way table

$$\begin{array}{ccc} \hline (1,1) & \cdots & (1,k_1) \\ \vdots & \ddots & \vdots \\ (j,1) & \cdots & \left(j,k_j\right) \\ \hline \end{array} \tag{B.1}$$

and let $v = \{v_1, \ldots, v_m\}$ be a partition of the elements of the table (B.1) into disjoint sets. We say that two sets v_l and v_k of the partition v hook if there exist $(j_1, i_1) \in v_l$ and $(j_2, i_2) \in v_k$ such that $j_1 = j_2$. We say that the sets v_l and v_k communicate if there exists a sequence of sets

$$v_{i_1} = v_l, v_{i_2}, \ldots, v_{i_m} = v_k$$

such that v_{i_j} and $v_{i_{j+1}}$ hook for each j. A partition is said to be indecomposable if all its sets communicate.

Lemma B.1 (*Leonov and Shiryaev [47]*). *Given an array $\{y_{mn}\}$, $m = 1, \ldots, j$, $n = 1, \ldots, k_m$, of random variables, we consider the complex-valued random variables*

$$z_m = \prod_{n=1}^{k_m} y_{mn}, \quad m = 1, \ldots, j.$$

Then the following formula for the joint cumulant of the variables z_m holds:

$$cum\left(z_1, \ldots, z_j\right) = \sum_v c_{v_1} \ldots c_{v_p}, \tag{B.2}$$

where the sum extends over all indecomposable partitions $v = (v_1, \ldots, v_p)$ of the table (B.1) *and*

$$c_{v_i} = cum\left(\{y_{lk}\}, (l, k) \in v_i\right).$$

For the proofs of the theorems of this paper, it will be more convenient to use another set of indices when we consider an array of the form $\{y_{mn}\}$, $m = 1, \ldots, j$, $n = 1, \ldots, k_m$, and apply the formula (B.2). Namely, instead of the double index (m, n), we use as index an integer denoting the position at which (m, n) appears in the list $\{(1, 1), \ldots, (1, k_1), (2, 1), \ldots, (2, k_2), \ldots, (j, 1), \ldots (j, k_j)\}$. That is, the table (B.1) can be written as

$$\begin{array}{ll} \hline 1 & \cdots k_1 \\ k_1+1 & \cdots k_1+k_2 \\ \vdots & \ddots \vdots \\ \sum_{i=1}^{j-1} k_i + 1 \cdots & \sum_{i=1}^{j} k_i \\ \hline \end{array}$$

As an application of Lemma B.1 we have

$$E\left(y_1 \cdot \ldots \cdot y_k\right) = \sum_v c_{v_1} \cdot \ldots \cdot c_{v_p}, \tag{B.3}$$

where the summation extends over all partitions $\left(v_1, \ldots, v_p\right)$ of the set $(1, \ldots, k)$ and c_v denotes the joint cumulant of the y's with indices in v.

We use also the following algebraic property of cumulants:

$$cum\left\{\sum_{i=1}^{n_1} a_{1i}y_{1i},\ldots,\sum_{i=1}^{n_j} a_{ji}y_{ji}\right\}=\sum_{i_1=1}^{n_1}\ldots\sum_{i_j=1}^{n_j} a_{1i_1}\ldots a_{ji_j}\,cum\left(y_{1i_1},\ldots,y_{ji_j}\right). \quad \text{(B.4)}$$

References

[1] V.G. Alekseev, Asymptotic properties of higher-order periodograms, Theory Probab. Appl. 40 (1995) 409–419.
[2] S.J. Anderson, A.R. Mohoney, A.O. Zollo, Applications of higher-order statistical signal processing to radar, in: B. Boashash, E.J. Powers, A.M. Zoubir (Eds.), Higher-Order Statistical Signal Processing, Longman, New York, 1995, pp. 405–446.
[3] V.V. Anh, C.C. Heyde, N.N. Leonenko, Dynamic models of long-memory processes driven by Lévy noise, J. Appl. Probab. 39 (4) (2002) 730–747.
[4] V.V. Anh, N.N. Leonenko, L.M. Sakhno, On a class of minimum contrast estimators for fractional stochastic processes and fields, J. Statist. Plann. Inference 123 (1) (2004) 161–185.
[5] V.V. Anh, N.N. Leonenko, L. Sakhno, Quasilikelihood-based higher-order spectral estimation of random fields with possible long-range dependence, J. Appl. Probab. 41A (2004) 35–53.
[6] V.V. Anh, N.N. Leonenko, L.M. Sakhno, Minimum contrast estimation of random processes based on information of the second and third orders, J. Statist. Plann. Inference, 2006, in press, doi:10.1016/j.jspi.2006.03.001.
[7] F. Avram, M. Taqqu, Applications of a generalized Brascamp–Lieb–Barthe inequality and a generalized Szegö theorem to the asymptotic theory of sums, integrals and quadratic forms, in: P. Bertail, P. Doukhan, P. Soulier (Eds.), Dependence in Probability and Statistics, Springer, New York, 2006.
[8] M. Baba Hara, Statistical estimation of higher order spectral densities by means of general typering, Appl. Math. 24 (1997) 357–381.
[9] O.E. Barndorff-Nielsen, N. Shephard, Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial econometrics, part 2, J. Roy. Statist. Soc. B 63 (2001) 167–241.
[10] O.E. Barndorff-Nielsen, N. Shephard, Econometric analysis of realized volatility and it use in estimating stochastic volatility models, J. Roy. Statist. Soc. B 64 (2002) 253–280.
[11] R. Bentkus, Asymptotic normality of an estimate of the spectral function, Litovsk. Mat. Sb. 12 (3) (1972) 5–18.
[12] R. Bentkus, On the error of the estimate of the spectral function of a stationary process, Liet. Mat. Rink. 12 (1) (1972) 55–71.
[13] R. Bentkus, Cumulants of estimates of the spectrum of a stationary sequence, Liet. Mat. Rink. 16 (4) (1976) 37–61.
[14] R. Bentkus, R.R. Malyukyavichyus, Statistical estimation of a multidimensional parameter of a spectral density, Lithuanian Math. J. 28 (1988) 115–126.
[15] R. Bentkus, R. Rutkauskas, On the asymptotics of the first two moments of second order spectral estimators, Litovsk. Mat. Sb. 13 (1) (1973) 29–45.
[16] D.R. Brillinger, Introduction to polyspectra, Ann. Math. Statist. 36 (1965) 1351–1374.
[17] D.R. Brillinger, Fourier inference: some methods for the analysis of array and non-Gaussian series data, Water Resources Bull. 21 (5) (1985) 743–756.
[18] D.R. Brillinger, Some uses of cumulants in wavelet analysis, Nonparametric Statist. 6 (1996) 93–114.
[19] D.R. Brillinger, M. Rosenblatt, Asymptotic theory of estimates of *k*th order spectra, in: B. Harris (Ed.), Spectral Analysis of Time Series, Wiley, New York, 1967, pp. 153–188.
[20] V. Buldygin, F. Utzet, V. Zaiats, Asymptotic normality of cross-correlogram estimates of the response function, Statist. Inference Stochastic Process. 7 (2004) 1–34.
[21] K.C. Chan, G.A. Karolyi, F.A. Longstaff, A.B. Sanders, An sample comparison of alternative models of the short-term interest rate, J. Finance 67 (1992) 1209–1227.
[22] J.C. Cox, J.E. Ingersoll, S.A. Ross, A theory of the term structure of interest rate, Econometrica 53 (2) (1985) 385–407.
[23] W. Dusmuir, E.J. Hannan, Vector linear time series models, Adv. Appl. Probab. 8 (1976) 339–360.
[24] R.F. Dwyer, Range and Doppler information from fourth-order spectra, IEEE J. Oceanic Eng. 16 (3) (1991) 233–243.
[25] R. Fox, M.S. Taqqu, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series, Ann. Statist. 14 (2) (1986) 517–532.
[26] J. Gao, V. Anh, C. Heyde, Statistical estimation of nonstationary Gaussian processes with long-range dependence and intermittency, Stochastic Process. Appl. 99 (2002) 295–321.

[27] J. Gao, V. Anh, C. Heyde, Q. Tieng, Parameter estimation of stochastic processes with long-range dependence and intermittency, J. Time Series Anal. 22 (5) (2001) 517–535.
[28] L. Giraitis, D. Surgailis, A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle estimate, Probab. Theory Related Fields 86 (1990) 87–104.
[29] L. Giraitis, M.S. Taqqu, Convergence of normalized quadratic forms, J. Statist. Plann. Inference 80 (1999) 15–35.
[30] G. Grimmett, Weak convergence using higher-order cumulants, J. Theoret. Probab. 5 (1992) 767–773.
[31] X. Guyon, Parameter estimation for a stationary process on a *d*-dimensional lattice, Biometrica 69 (1982) 95–105.
[32] E.J. Hannan, Multiple Time Series, Springer, New York, 1970.
[33] E.J. Hannan, The asymptotic theory of linear time series models, J. Appl. Probab. 10 (1973) 130–145.
[34] C. Heyde, R. Gay, Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence, Stochastic Process. Appl. 45 (1993) 169–182.
[35] C.C. Heyde, Quasi-Likelihood and its Application. General Theory of Optimal Parameter Estimation, Springer, New York, 1997.
[36] I.A. Ibragimov, On estimation of the spectral function of a stationary Gaussian process, Theory Probab. Appl. 8 (4) (1963) 366–401.
[37] I.A. Ibragimov, On maximum likelihood estimation of parameters of the spectral density of stationary time series, Theory Prob. Appl. 12 (1) (1967) 115–119.
[38] E. Igloi, G. Terdik, Bilinear stochastic systems with fractional Brownian motion input, Ann. Appl. Probab. 9 (1) (1999) 46–77.
[39] E. Iglói, G. Terdik, Superposition of diffusions with linear generator and its multifractal limit process, ESAIM: Probab. Statist. 7 (2003) 23–88.
[40] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, Amsterdam, 1981.
[41] I. Joury, Complex modulation and bispectral analysis of radar signals, in: B. Boashash, E.J. Powers, A.M. Zoubir (Eds.), Higher-Order Statistical Signal Processing, Longman, New York, 1995, pp. 447–476.
[42] G. Kallianpur, Stochastic Filtering Theory, Springer, New York, 1980.
[43] D.M. Keenan, Limiting behavior of functionals of higher-order sample cumulant spectra, Ann. Statist. 15 (1987) 134–151.
[44] M. Kumon, Identification of transfer function matrix using higher-order spectra, Japan J. Indust. Appl. Math. 13 (1996) 217–233.
[45] N.N. Leonenko, E.M. Moldavs'ka, Minimum contrast estimators of a parameter of the spectral density of continuous time random fields, Theory Probab. Math. Statist. 58 (1999) 101–112.
[46] N.N. Leonenko, A.Y. Sikorskii, G. Terdik, On spectral and bispectral estimator of the parameter of nongaussian data, Random Oper. Stochastic Equations 6 (2) (1998) 159–182 Correction: ibidem 7 (1999) 107.
[47] V.V. Leonov, A.N. Shiryaev, On a method of calculation of semi-invariants, Theory Probab. Appl. 4 (1959) 319–329.
[48] C.L. Nikias, A.P. Petropulu, Higher-Order Spectra Analysis. A Nonlinear Signal Processing Framework, Prentice-Hall, New Jersey, 1993.
[49] C. Pezeshki, V. Chandran, Using higher-order spectra for the analysis of chaotic systems, in: B. Boashash, E.J. Powers, A.M. Zoubir (Eds.), Higher-Order Statistical Signal Processing, Longman, New York, 1995, pp. 477–490.
[50] M. Rosenblatt, Stationary Sequences and Random Fields, Birkhäuser, Boston, 1985.
[51] R.J. Serfling, Approximation Theorems of Mathematical Statistics, Wiley, New York, 1980.
[52] B. Sundt, J. Dhaene, N. De Pril, Some results on moments and cumulants, Scand. Actuarial J. 1 (1998) 24–40.
[53] A. Swami, G.B. Giannakis, G. Zhou, Bibliography on higher-order statistics, Signal Process. 60 (1997) 65–126.
[54] G. Terdik, Bilinear stochastic Models and Related Problems of Nonlinear Time Series Analysis, Lecture Notes in Statistics, vol. 142, Springer, New York, 1999.
[55] T. Yamada, S. Watanabe, On the uniqueness of solutions of stochastic differential equations, J. Math. Kyoto Univ. 11 (1971) 155–167.
[56] H.-C. Zhang, P. Shaman, On the calculation of cumulants of estimator arising from a linear time series regression model, J. Multivariate Anal. 37 (1991) 135–150.